

Opportunity in Kyoto for a president's legacy

The success or failure of global climate negotiations rests on the shoulders of Bill Clinton, who can enhance his stature by holding his nerve in the face of industry's scare-mongering.

After more than ten years of heated scientific and political argument about the threat of anthropogenic climate change, the decisive moment approaches. December's conference of nations in Kyoto will either commit the planet to a sensible response, or visibly fail to do so. If there is one person who can singlehandedly have a major impact on the outcome, it is the president of the United States.

This week in Washington, President Bill Clinton gathered together an impressive array of scientific and political talent to discuss publicly the global warming issue. The event was the culmination of a serious, if somewhat belated, attempt by the president to offer leadership to the American people on this issue.

Although the signals coming from Clinton's advisers are mixed and inconclusive, the president's own recent statements suggest that he is determined that there should be a meaningful agreement at Kyoto. Last week, in an imaginative public relations initiative, he invited television weathercasters (many of whom conceal advanced meteorological degrees underneath their cheerful countenances) from all over the United States to the White House for an education session on global warming. "I want to try to get America to accept the fact that the majority scientific opinion, the overwhelming majority scientific opinion, is accurate," he told them.

What is required now is for the administration to muster the political courage it regrettably needs, to commit itself to firm targets and timetables for emission reductions that it would like to achieve, and to go to Kyoto in good faith to negotiate a deal. But that is easier said than done. The United States was due to publish its targets in June: so far it has failed to do so, causing much exasperation elsewhere. But inside the United States, that failure is seen as a sensible precaution against a lengthy onslaught from the industrial interests that oppose any meaningful agreement.

The debate in each developed country has been substantially similar to that in the United States. American aversion to taxation — and energy taxation in particular — has clearly played a role in that country's laggardly approach to climate change. But Americans are not alone in disliking taxes. US public awareness of environmental issues in general is high. But the lack of a firm political culture allows industrial and labour interests the opportunity to buy influence in the debate. From the outset, those interests have sought to nurture a group of scientific nay-sayers, many from outside the field of climate science. This group has distorted the scientific debate, perpetuating a misleading argument about the absence of proof of an anthropogenic component to climate change at a time when the principal argument in the scientific literature was about the scale of that component.

That strategy now appears to be coming apart. Opinion polls suggest that it has had little impact on US public opinion (see page 531). More tellingly, industry has switched the focus of its attack: in its latest, \$15-million television advertising campaign, xenophobia has replaced junk science as the central argument for inaction. The decision of Ted Turner's CNN network to stop running the advertisements is an imperious declaration of good taste.

The effectiveness of industry's direct overtures to the public will fortunately be limited by its past record of crying wolf on the threats posed by previous environmental measures. The same groups vigorously and unsuccessfully opposed clean air, clean water, catalytic converters and restrictions on sulphur dioxide emissions and CFCs, to name but a few. But these industrial and labour interests are, one way and another, able to exert strong influence on the 100 senators who must ratify any agreement to come out of Kyoto. The Senate has already exercised its constitutional right to advise the president on the negotiations by means of a non-binding resolution, sponsored by Robert Byrd (Democrat, West Virginia), which it passed unanimously in August. The Byrd resolution contains potential obstacles to agreement in Kyoto, but these are not insurmountable. It asks that any agreement "mandates new specific scheduled commitments to limit or reduce greenhouse gas emissions for developing country parties within the same compliance period" as the agreement between developed countries.

Proposals

US officials do not expect developing countries to join in any agreement reached by developed countries at Kyoto: they hope to persuade them to accept an appendix to the agreement, which might, for example, set out a binding timetable for negotiations that will bring the developing countries into the process at a later date. Treaty opponents want Clinton to push this point hard, perhaps hoping to prompt a breakdown in the talks: they will surely be disappointed.

The other issue that America will bring to Kyoto is a proposal for the trading of pollution permits between governments, to ease compliance. That would allow rich countries to buy their way out of trouble through bilateral agreements with poorer countries. After some early suspicion, other countries are coming round to the idea, provided that the permits are for deliberate future cuts in emissions, and not for past accidental events, such as the slump in emissions from Eastern Europe and Russia in the early 1990s and recent declines in UK emissions following privatization of energy suppliers.

Evidence of good faith from the developing countries, together with acceptance of permit trading, could open the way for an historic agreement at Kyoto. Such an agreement may require the United States to go further than the stabilization proposal that the Clinton administration is now considering. Japan's target of an average 5 per cent cut below 1990 levels by 2010 is the minimum that developing countries should demand as evidence of the developed world's own good intentions. Europe's role will be as a broker between the United States and the developing countries. Australia, which is opposing strict limitations, can expect to find itself thoroughly isolated.

Bringing back such an agreement and selling it will require considerable courage from a Democrat administration that can expect to be attacked where it hurts — on the tax issue. At last, Clinton seems to be warming to a task befitting a flawed but popular president in search of a legacy. Success at Kyoto requires no less. □



Nobel panel rewards prion theory after years of heated debate

[LONDON] The 1997 Nobel Prize in Physiology and Medicine was awarded this week to Stanley Prusiner of the University of California, San Francisco, for his contributions towards the identification of the infectious agent that causes transmissible spongiform encephalopathies (TSEs). This class of disease includes BSE in cows, scrapie in sheep and Creutzfeldt–Jakob disease and its new variant in humans.

The award has been greeted with delight for Prusiner by scientists across the world. Charles Weissmann of Zurich University says he is convinced that Prusiner “deserves the prize” and is “very pleased”, while Colin Masters of the Walter and Eliza Hall Institute in Melbourne, Australia, describes Prusiner’s work as “terrific”.

But the award has also generated some surprise, as Prusiner’s ideas remain unproven. It is widely agreed that TSEs are most probably caused by a completely new type of agent that does not contain nucleic acids. But the infectious agent remains unknown, and there is some concern in the scientific community that the Nobel Assembly may be prematurely endorsing what remains a hypothesis, however appealing.

It was 21 years ago — at the beginning of Prusiner’s research career — that Carleton Gajdusek was made a Nobel laureate for demonstrating that human spongiform encephalopathies can be transmissible diseases. Since then, Prusiner has been one of the leaders of attempts to determine the agent which is responsible for this infectivity, and which he claims is an abnormal form of the host PrP protein. Somehow the disease form of this protein is able to catalyse its own formation by converting its normal, cellular homologue into the disease form.

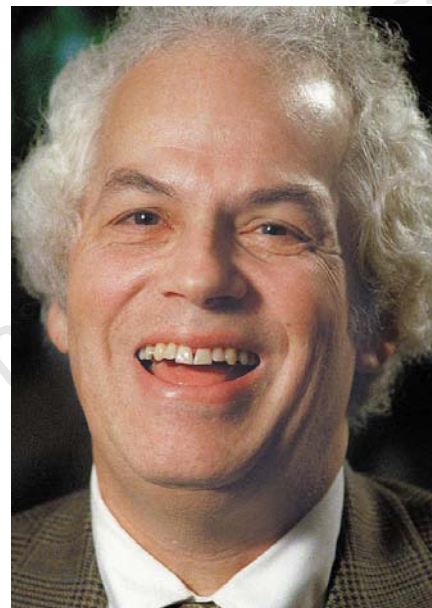
This hypothesis, known as the protein-only hypothesis, implies a simple disease agent — which Prusiner called the ‘prion’, a rather tortured acronym of ‘small proteinaceous infectious particle’ — distinct from all known disease agents. The difference is that the prion probably does not contain nucleic acid, unlike bacteria, viruses, fungi and multicellular parasites.

In the 1980s Prusiner championed the protein-only hypothesis in the face of hostile criticism, at a time when many scientists were convinced that all life-forms needed nucleic acid to replicate. Those close to the debate at the time say that the severity of the criticism directed at Prusiner was extraordinary, and obviously had an impact on him.

The idea of a proteinaceous agent for TSEs was first suggested in 1966 by Tikvah Alper, working for the Medical Research Council at the Hammersmith Hospital in London, who found that ultraviolet radiation that destroys all nucleic acid does not destroy scrapie infectivity (see *Nature* 214, 764–766; 1966).

The following year, J. S. Griffith at Bedford College, London, proposed that infectivity in scrapie could be due to the altered conformation of a normal cellular protein (see *Nature* 215, 1043–1044). Almost a decade later, Prusiner developed these ideas by producing compelling biochemical analyses to confirm his views and defended them rigorously against strong criticisms.

Prusiner’s career has involved many productive collaborations with scientists who have been inspired by his exciting theory and the strength of his convictions. He purified the PrP protein from infectious material and showed that it was a major component of the prion, and then sequenced the protein.



Stanley Prusiner, pictured above on Monday in Bethesda, Maryland, said he felt he had been ‘vindicated’ by his award of the Nobel prize.

This led to the cloning of the PrP protein by Bruno Oesch and Konrad Basler under the direction of Charles Weissmann in 1985. Weissmann is recognised as having made particularly important conceptual and molecular biology contributions to the “protein-only” hypothesis and the course of prion research over the last 20 years.

If conclusively proven, it is widely agreed that the protein-only hypothesis and Prusiner are fully deserving of a Nobel prize. But scientific issues remain unresolved. Although it is possible to get the disease form of PrP to convert to normal PrP in a test tube (see *Nature* 375, 698–700; 1995) the resultant material is not infectious.

Perhaps most significantly, there are many strains of TSEs, each with their own characteristics (much like strains of flu). In fact there are too many strains of TSEs to be easily explained by protein conformation alone. Finally, Prusiner himself claims that there is likely to be an as yet unidentified protein accomplice to PrP, which he calls protein X, needed for infectivity.

But Prusiner has certainly been an extraordinary motivator in the field of TSEs over the last 20 years. Indeed, perhaps it is the most productive research that raises as many questions as it answers. Peter Lansbury, a chemist at Harvard University, says that Prusiner “took a medical and biological mystery and changed it into a biophysical and biochemical issue with a very original idea. And that’s fabulous”.

Harriet Coles

Japan agrees on target for climate curbs

[TOKYO] A very public tussle between Japan’s trade and environment ministries came to an end this week with the belated announcement that the country will support a worldwide average target of reducing greenhouse gas emissions by 5 per cent from 1990 levels between 2008 and 2012.

For Japan itself the proposal amounts to a reduction in emissions of just 2.5 per cent. The final figure was rounded on by environmentalist groups as too little, too late. It is also likely to create disappointment in European capitals, which had hoped for more from the host of this year’s annual conference of the United Nations climate

convention in December.

But Japan’s Ministry of International Trade and Industry argued until the last minute that commitments to reduce emissions below 1990 levels would hurt Japanese industry. And the United States may take comfort from the announcement.

Japan has argued for a flexible system of greenhouse cuts that imposes fewer obligations on countries that promote energy efficiency and curb population growth. Under its formula, Japan’s record of energy efficiency and low population growth entitles the country to halve its 5 per cent target. (See also pages 531 and 532.) □

Congress warns NASA not to raid science funds

[WASHINGTON] The US Congress told the National Aeronautics and Space Administration (NASA) last week that it must solve its cost problems in building the international space station without raiding funds allocated to science programmes.

House of Representatives and Senate appropriations committees granted the agency \$148 million beyond its \$13.5 billion request for the 1998 fiscal year, which began last week. But they specifically denied a request from NASA to transfer money from the science, aeronautics and technology account to help pay for a projected shortfall of \$430 million in 1998 for space station construction.

While sending a strong signal that Congress will not allow the station to cannibalize the agency's Earth and space science programmes, the move puts NASA in a difficult position. Even with an extra \$100 million for the station, and permission to move \$130 million in "mission support" (largely salaries and administrative costs) and other funds, the agency remains \$200 million short of what it says it needs in the coming year to keep the station on track.

The agency's administrator, Daniel Goldin, warned a Senate committee last month that without \$430 million in new funding and transfer authority, work on the station could slip. This in turn could delay its 2003 completion date and increase its total cost of \$17 billion.

NASA and its main contractor for the station, Boeing, estimate the station will be \$600 million over budget by the time it is finished. Only \$100 million of the 1998 shortfall is caused by problems in accommodating Russian hardware delays, the rest being attributed to schedule overruns and development problems at Boeing.

Congress has become increasingly alarmed at growing projections for the shortfall, and is therefore imposing strict conditions on NASA's 1998 appropriation. The agency will receive only about two-thirds of the \$2.35 billion allocated to the station until it produces (by 31 March) accounts showing exactly how much the project will cost to complete.

NASA will be asked to cooperate in this exercise with the congressional General Accounting Office (GAO), which has warned repeatedly of cost and schedule problems with the station, and has consistently estimated a higher final cost than NASA has.

Last week's decision showed once again the influence of Barbara Mikulski, the top Democrat on the Senate appropriations subcommittee that oversees NASA. Mikulski

supports the station, but is even more concerned that it should not jeopardize funding for the Mission to Planet Earth and space science programmes that make up the majority of work conducted at the Goddard Space Flight Center in her home state of Maryland.

The station funding crisis adds to growing confusion at the space agency about its budget. In August the White House Office of Management and Budget (OMB) surprised NASA managers by telling them to request only \$12.6 billion from Congress in 1999. That is a sharp reduction from the \$13.4 billion that NASA budget planners had been expecting, based on a five-year projection accompanying the administration's 1998 request.

Although NASA — and other federal agencies that also received low 1999 targets from OMB — have been told that their final request is likely to be higher, programme managers are left uncertain how much money they have to work with in the coming years.

If NASA had to live with a \$12.6 billion budget for 1999, says one agency official, it would be "disaster" and is likely to lead to cancellation of several high-priority proposed science initiatives. These include a robotic landing on Mars in 2001, a Mars sample return mission in 2005, a space-based interferometry mission, the LightSAR imaging radar for Earth observation, and proposed increases for basic research and technology development.

Tony Reichhardt

UK's greenhouse reduction target 'hard but feasible'

[LONDON] The British government's chief scientist, Sir Robert May, has acknowledged that the government's promise to reduce greenhouse gas emissions by 20 per cent of 1990 levels by 2010 will be "difficult" to achieve, but says it should remain a target.

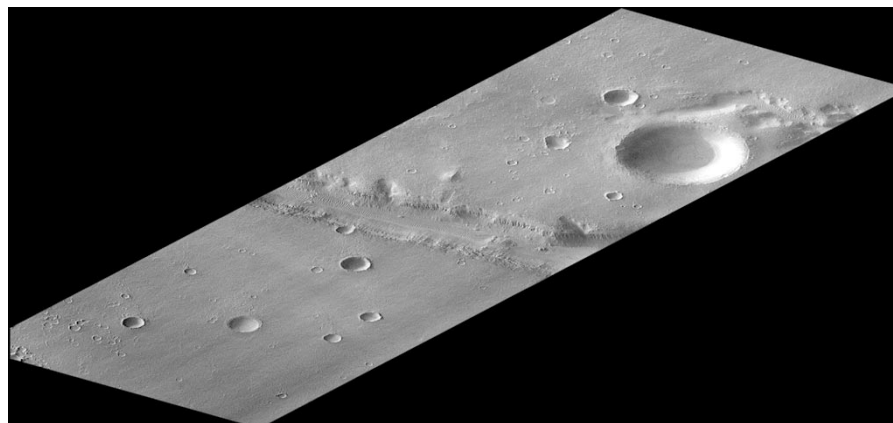
Speaking at the launch of a briefing document on climate change written for Prime Minister Tony Blair, May said that reductions would have to come from all sectors of energy use: road transport, domestic use and industry. "Industry should see climate change as opportunity not as threat," he said.

But it appears unlikely that the government will formally commit itself to a 20 per cent reduction if — as is expected — the December conference of the United Nations climate convention in Kyoto, Japan, agrees on lower greenhouse gas targets. The UK government has said it will not outline specific policies and measures until well after the Kyoto conference.

An interdepartmental working group of officials from the departments of the environment and trade, and the Treasury, has begun to look into how Britain could achieve a 20 per cent reduction. An environment department spokesman says that the group is not expected to report until the middle of 1998 at the earliest. Meanwhile, the combined Department of the Environment, Transport and the Regions has begun drawing up proposals for an integrated transport policy.

Ehsan Masood

Surveyor maps desert valleys of Mars



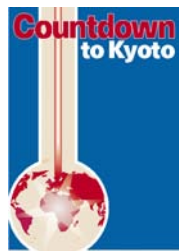
[LONDON] Preliminary images sent back by the orbiting Mars Global Surveyor show a desert-like terrain pock-marked with craters. The mapped region (above) is called Nirgal Vallis, which is one of a number of canyons known as valley networks.

The origin of these valleys is the subject

of debate. Some believe they were formed by water flowing across the surface, but other suggest they were formed by the collapse and erosion associated with groundwater processes.

A closer look at the terrain is likely to help resolve the controversy.

US target 'likely to equal 1990 emissions'



[WASHINGTON] US President Bill Clinton reiterated his support this week for action to reduce greenhouse gas emissions. But the proposal he appears set to unveil for December's Kyoto climate conference in Japan will disappoint environmentalists and European nations by seeking to cut emissions back only to their 1990 levels by 2010, at the earliest, and not below these levels, as others are pushing for.

Clinton told a special White House conference at Georgetown University in Washington DC on Monday that the United States would deal with the threat of global warming by taking "a moderate and disciplined approach" to reducing emissions. "In Kyoto, we will ask for a meaningful, but equitable, response from all nations," he said.

The president indicated that he sides with those seeking decisive action to cut emissions. "We do not want to burden our children with our failure to act," he declared. But he also came close to ruling out direct taxation on fuel — a 'carbon tax' — as a means to that end. "If all we do is raise the price of fuel we are not going to get what we want. It won't pass the Senate and it won't pass muster with the American people," the president said.

The US Senate must ratify any agreement reached in Kyoto, and Clinton prefers other options, such as reducing the price of 'clean' energy, introducing a permit system for industrial users of fossil fuels, and improving consumer awareness of the implications of their own energy choices.

Monday's day-long meeting, attended by most of Clinton's cabinet members, along with top economists, climate scientists and business leaders, was an attempt to raise public awareness of the global warming issue in the run-up to the Kyoto meeting.

At least one leading environmental lobbyist in Washington believes the set of policy options now being considered by Clinton's advisers would merely reduce greenhouse gas emissions to 1990 levels by 2010, by 2020 or even by 2030.

Options that would reduce emissions below 1990 levels by 2010 have been discounted, says lobbyist Phil Clapp, a view confirmed by several environmentalists and scientific advisers to the government, although one claimed there was still time to revive these options before a position is announced.

Administration officials say that the announcement will take place at or just before a pre-Kyoto negotiating session, due to take place in Bonn on 20 October. "I'm hoping to have a target early in the day at

Bonn," Melinda Kimble, acting assistant secretary of state, said last week.

As part of its position, the United States will ask developing countries to place limits on their greenhouse gas emissions at a future date. "One way or another, the developing countries must be making some kind of commitment," says Bruce Babbitt, the interior secretary. "But only people who want to kill the process say that every [country] should do the same thing."

At the Georgetown conference, Al Gore, the vice-president, questioned the moral high ground that Europe, which is seeking agreement on a 15 per cent reduction by 2010, is claiming in the climate treaty negotiations. He said that Germany had benefited from the closure of polluting industries in the former East Germany after 1990, and that the United Kingdom had cut emissions by shutting down most of its coal industry and switching to gas at that time. "It is kind of a freebie for them," the vice-president said.

Economists continue to disagree sharply about the costs of action to cut emissions in the United States. The US is now emitting about 12 per cent more carbon dioxide than it did in 1990, and it is widely accepted that if it took no action, emissions would grow by 25 to 30 per cent above 1990 levels by 2010.

Paul Krugman, professor of economics at

the Massachusetts Institute of Technology, says that the impact of a move to stabilize emissions would be minimal. "On the scale of economic dislocations we have in this country, this is not a major one," he told a meeting of the Union of Concerned Scientists last week. "Unless we completely mishandle this, it'll be lost in the noise."

But Bill Nordhaus, professor of economics at Yale and the most sceptical of the 20 experts on panels at Clinton's Georgetown conference, said that stabilization could involve doubling the cost of fuel, and would be "a bigger deal than the oil price shocks of the 1970s".

A scheme for trading permits that allow emissions of carbon dioxide is likely to play a major role in Clinton's plan to implement anything agreed in Kyoto. The administration is encouraged by the success of a permit scheme in reducing emissions of sulphur dioxide to stop acid rain.

"The potential for carbon dioxide reduction is huge," said Tom Casten, president of Trigen, an energy conservation company.

John Adams, executive director of the Natural Resources Defence Council, an environmental group, said afterwards: "You had to come away with a sense of the deep commitment that Clinton has. Now, it's his ballgame to win or lose."

Colin Macilwain

Public rejects sceptics' line on global warming

[WASHINGTON] Two-thirds of US voters regard global warming as a serious threat and support an international commitment to cut greenhouse gas emissions, even if this leads to higher energy prices, according to a public opinion poll for the World Wildlife Fund.

The poll, which was conducted in August before a recent barrage of industry-sponsored television advertisements attacked such a treaty, also found that voters believe the scientific debate about global warming has been settled.

The result suggests that a longer-running campaign by the Global Climate Coalition, an industrial pressure group, to undermine public confidence in the mainstream view of climate scientists about manmade global warming has made little headway with

the general public.

The campaign has provided a platform for global warming sceptics. But 75 per cent of those polled agreed with a statement that "the only scientists who do not believe global warming is happening are paid by big oil, coal and gas companies to find the results that will protect business interests".

Only 15 per cent accepted the opposing view, that "scientists disagree among themselves" and "there is no real evidence that carbon dioxide emissions from coal, oil and gasoline are causing global warming".

Intriguingly, among the subset of the polling sample who actually work in the energy or automobile industries, scepticism about industry's position ran even stronger: 79 per cent accepted the first statement,

complete with its withering criticism of their own industry.

The poll is likely to encourage environmentalists inside the Clinton administration, who are trying to persuade the president that a strong agreement to cut greenhouse gas emissions at the forthcoming Kyoto meeting will be politically popular.

But the telephone survey, conducted among 800 registered voters by the Mellman Group for the World Wildlife Fund, gave a mixed message on what Americans were ready to pay for such a treaty. Of the sample, 58 per cent opposed a 50 cent increase in the gasoline tax — gasoline costs \$1.20 a US gallon. But 61 per cent said they were prepared to pay \$20 a month to buy cleaner energy from their electricity supplier.

C.M.

Cash-strapped ANZAAS is set for closure

[ADELAIDE] The 109-year-old Australian and New Zealand Association for the Advancement of Science (ANZAAS) announced last week that it lacks sufficient funds to continue operating. The association has been beset by falling membership and declining attendance at its congresses.

The association sustained a heavy loss on its 1996 congress in Canberra, and last week's congress in Adelaide, South Australia, only broke even. Chairman Bruce McKellar, dean of science at the University of Melbourne, is to recommend to a special general meeting in December that it be wound up.

Created as an offshoot of the British Association for the Advancement of Science in 1888 when Australia was a colony, ANZAAS gave coherence and identity to the emerging scientific communities both of Australia and of New Zealand, which hosted a congress every decade.

Congresses attracted the largest gatherings of scientists in Australia and New Zealand, with a wide range of specialized sections and, uniquely in the scientific calendar, crossdisciplinary sessions and forums on policy and social issues. But after the Melbourne congress in 1985 and the last in New Zealand in 1987, attendances fell sharply.

One factor has been a rapid growth in specialist conferences held at the same time. Last week, for example, the Australian Society for Microbiology also met in Adelaide, drawing 900 paying delegates. By contrast, sessions at ANZAAS had audiences of 10 to 20. The general registration of 220 indicated



Beginning of the end? Chairman McKellar (left) plans to ask for ANZAAS to be wound up.

no resurgence of interest and finance in the long term.

Local organizers had had to use volunteer labour, as the federal government initially refused to provide the congress's normal grant. Under pressure, Peter McGauran, the science minister, relented. But his offer of A\$14,000 (US\$10,000) was too small and too late to make a difference.

The turning point came when the ANZAAS council suddenly had to cancel the 1998 meeting in Hobart, Tasmania, just as it was due to be publicized at Adelaide. Con-

venor Jim Reid of the University of Tasmania says that it withdrew in the face of competition for government funds and corporate sponsorship. Another factor was a challenge from Australian Science Communicators, a media group dissatisfied with ANZAAS.

With no congress next year, ANZAAS's position became hopeless. The viability of its official journal, *Search*, Australia's only science magazine, is also threatened and a rescue campaign has begun, supported by Sir Gustav Nossal, president of the Australian Academy of Science.

Some link the demise of ANZAAS with McGauran's forced resignation in the same week (see *Nature* 389, 427; 1997) as indicating malaise in the public standing of science in Australia. But Nossal hopes that a new organization, representing individual scientists, may rise from the ashes of ANZAAS.

McKellar is recommending that the organization pass its remaining funds of A\$100,000 to organizations with "similar aims" in the promotion and public understanding of science.

Peter Pockley

Cabinet minister takes over Australian science portfolio

[ADELAIDE] Prime Minister John Howard announced last weekend that John Moore, the cabinet minister in the Ministry for Industry, Science and Tourism, will take on the responsibilities of Peter McGauran, who resigned last

week as science minister (see above).

In a reshuffle which was prompted by three ministerial resignations, controversial education minister Amanda Vanstone has been replaced by the former minister for

schools, David Kemp. A former professor of political science, Kemp is expected to show himself more knowledgeable about the workings of universities. He is considered to be an economic hardliner. □

Plagiarism claims turn spotlight on US Navy's misconduct policy

[SAN DIEGO] A psychologist alleged to have plagiarized at least three books has been picked to head the research office at a major US Navy facility. The appointment raises questions about how the Navy deals with cases of scientific misconduct.

Dennis L. Reeves, who is 46, was named in August as head of the Clinical Investigation Department at the Naval Medical Center in San Diego, California, which monitors a broad array of research at several institutions in the region.

The promotion occurred at a time when the state of Maryland, where Reeves received his psychology licence, was considering taking disciplinary action because of the long-simmering charges of plagiarism. A psychologist licensed in one state can practise at any Navy facility.

Reeves has apologized to some of the authors in letters for what he termed inadvertent "borrowing" of material that was not "adequately attributed". Reeves and his superiors declined to be interviewed.

Like many US institutions, the Navy has no specific policy on probes into scientific misconduct. Edie Rosenthal, a Navy spokeswoman, says there are no plans to develop such a policy, because "our judicial system works well" on such matters.

The Navy examined some aspects of the issue in 1994 when the allegations first surfaced, deciding then that there was no plagiarism. But Navy records and interviews indicate that the inquiry was limited in scope. A colleague of Reeves examined the allegation, deciding in November 1994 there was only "editorial and style error".

But the Navy apparently did not scrutinize the remainder of the book, *The Clinical Assessment of Memory: A Practical Guide*, by Reeves and Danny Wedding, a psychologist at the University of Missouri who directs the Missouri Institute of Mental Health in St Louis.

Shortly after it was published by Springer Publishing Co. of New York, a nationally renowned psychologist, Muriel D.

Lezak, claimed that the book included sections from one of her books published by Oxford University Press (OUP).

After demands from OUP, Springer ceased distribution of the Reeves/Wedding book in August 1995, destroyed the remaining copies and paid Lezak a \$2,000 settlement. "Plagiarism is very high on my list of disgusting acts," commented Lezak, of the Oregon Health Sciences University.

She says she later discovered that parts of the Reeves/Wedding book included passages from psychologists at Harvard University, the Mayo Clinic in Minnesota and the University of California at San Diego.

Several of these authors complained to OUP. They also complained to the American Psychological Association (APA) about Wedding, who, unlike Reeves, is an APA member. Both Wedding and the APA refused to say what action, if any, was taken.

Although credited as co-author, Wedding says he wrote only two chapters, and neither included plagiarized material.

Reck Dalton

French labs funded by 'big science' cuts

[PARIS] Claude Allègre, the French minister for national education, research and technology, this week confirmed a large increase in research jobs, launched the country's first scheme of postdoctoral fellowships, and promised to reorganize France's research agencies.

In his first major policy statement on research since taking office in June, Allègre also cut funding for 'big science' facilities, and said he intended to end France's love affair with manned space flight, focusing its space activities on scientific and commercial goals. And he announced plans to set up a personal advisory board, half of whose members will be drawn from other European countries.

The FF53 billion (US\$8.9 billion) civil budget for research and development, which was announced last week (*Nature* 389, 432; 1997), will create 600 posts for scientists in the research organizations, 121 of them available immediately. Combined with 800 new posts in universities, as well as posts left vacant through retirement, there will be about 3,500 employment opportunities for young scientists this year.

The increase in scientific posts has been widely welcomed. After a fall in recruitment in recent years, "it is incontestably an important reversal of the past trend," says Henry-Edouard Audier, a chemist at the Ecole Polytechnique and a member of the board of the National Union of Scientific Researchers (SNCS).

Recruitment is the main preoccupation of the union, which argues that it is urgently needed to maintain a sensible age structure in the country's scientific workforce, with

half scheduled to retire over the next decade.

A new postdoctoral scheme has been allocated FF50 million. This will initially fund 250 posts in industry. But the government intends to extend the scheme to research organizations and universities after discussions with the organizations and trade unions. A special venture capital fund will also help postgraduates to create their own companies.

Allègre also promised on Monday (6 October) to seek an end to research tax credits for large companies — "they will have to invest themselves" — and to concentrate such credits on small and medium-sized companies, which are less able to afford research. Ministry officials say that credits for companies will be linked directly to their targets for recruiting scientists.

The research ministry itself is being reorganized to increase the emphasis on technological innovation. The ministry at present has 16 offices under its department for research and technology. It will now be split into two departments, one for research and the other for technology; Allègre says that the heavy day-to-day demands of research have in the past resulted in a neglect of technology. The number of offices will be reduced to 10, and women will be appointed to head up to half of them — these posts are currently all held by men.

Though many of these promises on employment have been welcomed, the unions are less happy about the funding for laboratories. This will increase this year from 2 to 3 per cent in research organizations, and by 5.4 per cent in universities.



Allègre: creating 3,500 posts for young scientists but cutting funding for space research.

Allègre has already promised to shift the emphasis from large programmes to investigator-driven research. But despite announcing that he will transfer FF300 million to achieve this from the FF4 billion budget for 'big science', the unions remain dissatisfied. "He has stabilized the situation, but there is still an effort to be made," says Audier, who points out that laboratory funds have been particularly hard hit by cuts in recent years.

The FF300 million reduction will mainly fall on microgravity research and other space programmes, although observers say that construction of the Virgo gravitational wave detector and Soleil synchrotron may also be delayed. Allègre this week reaffirmed his opposition to manned space flight, saying that France's priorities will be telecommunications, Earth observation, military satellites and robotic planetary missions. He recently appointed Gérard Brachet as director of CNES, the French space agency, with instructions to implement such reforms.

Allègre also said that France would respect its international commitments. But one ministry official says that France and Germany may negotiate spending ceilings on their contribution to the international space station, and pull out if these are exceeded. Allègre claimed that the United States wanted European countries to continue financing the station "to prevent us having our own independent programmes".

Other developments announced by Allègre include a new body to coordinate life science research among research organizations, and a reform of INSERM, the national biomedical research organization. A broader reform which would see a reduction in the number of research agencies is also on the cards, according to one ministry official, who says that, for example, life sciences research may be transferred out of the Atomic Energy Commission (CEA). **Declan Butler**

German research agencies feel the pinch

[MUNICH] Germany's two main scientific organizations, the Max-Planck-Society (MPS) and the Deutsche Forschungsgemeinschaft (DFG), which is Germany's university grants council, are likely to have their anticipated increase in federal funding cut from 5 per cent to about 3.5 per cent next year, as a result of a decision last week by the parliamentary commission on the budget.

Overall, Germany's 1998 federal budget for research and education would be trimmed by DM22 million (US\$12.5 million), to just

under DM15 billion. The parliamentary committee adjusted cabinet proposals for the federal budget for the MPS and the DFG, published last July, to DM785 million (down DM8 million) and DM1,070 million (down DM11 million) respectively.

In the past five years, the budgets of both organizations enjoyed a steady increase of five per cent a year, enabling them to catch up with costs from the reunification of German science. In July, Jürgen Rüttgers, Germany's research minister, promised a similar increase for 1998.

Germany's federal budget will be finally approved by parliament in late November, and the recommendations of the budget committee are usually accepted. But, before that, relatively minor adjustments may be made as a result of high-level negotiations. Hubert Markl and Wolfgang Frühwald, the presidents of the MPS and the DFG, have already reacted sharply to the proposed cuts. At a meeting in Jena last week, they asked Rüttgers for support and reminded him of his five per cent promise. **Quirin Schiermeier**

NCI apologizes for fallout study delay

[WASHINGTON] The director of the US National Cancer Institute (NCI) admitted last week that his institute had been wrong to delay releasing the preliminary findings of a study of Americans' exposure to radioactive iodine from above-ground nuclear tests in the 1950s.

Under sharp questioning from the Senate appropriations subcommittee on Labor, Health and Human Services and Education, Richard Klausner said "a more clear, more rapid and more aggressive plan for dissemination of the results to the public was called for".

The "essential nature" of the study was complete three to four years ago, and at that point "there should have been a plan put in place to publicly disseminate the information", Klausner said. He became head of the NCI in 1995.

Congress in 1982 mandated the NCI to investigate the exposure of Americans to radioactive iodine during 90 above-ground nuclear tests at the Nevada Test Site in the

1950s. But it was only in August, under pressure from critics and the media, that the institute released a summary of its findings.

These suggested that up to 75,000 additional thyroid cancers may result from the tests, mainly in those who were young children at the time. Perhaps 70 per cent have yet to be diagnosed. The full report, running to more than 100,000 pages, was issued last week on the Internet ('what's new' at <http://www.nci.nih.gov> or <http://rex.nci.nih.gov>).

The study did not address whether there is a definite link between thyroid cancer and internal exposure to radioactive iodine, mainly through ingestion of it in milk from cows. Instead it sought to establish the amount of radioactive iodine to which Americans on average were exposed. But Klausner said at last week's hearing that it is "very likely" that such exposure increases the risk of thyroid cancer.

Senator Tom Harkin (Democrat, Iowa), whose brother recently died of thyroid

cancer, who himself has thyroid nodules and who comes from an area that was highly exposed to the fallout, demanded to know why the study had taken almost 15 years to complete.

Klausner said that the study's length "overwhelmingly" reflected its complexity, and pointed to interim reports to Congress in 1984, 1986 and 1991, as well as open public meetings of an external advisory panel, as evidence that the institute was not trying to cover anything up.

Revealingly, a written chronology of the study's progress distributed by the NCI included comments made in 1993 such as "major part of database unintelligible to all but programmer" and "mapping software deleted after computer crash". It also said that the average dose to Americans from all 90 tests — 2 rads — was established in 1994.

Critics said NCI was negligent in not releasing those results immediately. "Once again, people feel like nuclear age guinea pigs or mere statistics — as though the government has infinitely more interest in studying them than in helping them," said Tim Connor, associate director of the Energy Research Foundation, who testified at the hearing.

Another critic, Joseph Lyon, a professor of family medicine at the University of Utah, suggested he had been effectively blocked by Bruce Wachholz, the study's overseer and chief of NCI's Radiation Effects Branch, from obtaining funding to follow up a study that found a threefold excess of thyroid cancer in subjects exposed as children to fallout from the tests. "The message came back: this is of no scientific interest," said Lyon.

Klausner denied that the NCI was not interested in follow-up studies, and said the denial of funding would have been the result of a failure to meet required standards of peer review. But Arlen Specter (Republican, Pennsylvania), the subcommittee chairman, said that the panel intended to pursue the matter. "These are very serious charges about this administration of NCI being... not up to doing this job," he said.

The NCI has now commissioned the Institute of Medicine to produce two reports determining the soundness of the NCI study's dose estimates, assigning risks of health consequences, and recommending public health measures. Both are due in the first half of 1998.

But their conclusions may come too late to avoid the damage to NCI's reputation. "The ground that's lost when, fairly or unfairly, a federal agency is perceived to be withholding information can be devastating," said Ruth Faden, director of the Johns Hopkins Bioethics Institute, who chaired the President's Advisory Committee on Human Radiation Experiments. **Meredith Wadman**

Gene-modified crop sabotaged in Ireland

[DUBLIN] Ireland's first genetically modified crop has been sabotaged by a group of environmental activists styling themselves the Gaelic Earth Liberation Front. Police launched a criminal investigation into the destruction of a one-acre crop of genetically modified sugar beet, being grown under licence by US chemical company Monsanto on a state research farm in County Carlow, about 50 miles from Dublin.

The crop had been genetically modified to resist the company's herbicide Roundup, and a three-year trial had been approved by Ireland's Environmental Protection Agency, which insists there is no threat to the environment. But, despite the assurance, the project caused a storm of protest, with environmentalists staging demonstrations, mounting pickets and taking court action in an effort to stop it.

Six months ago, the Irish High Court finally gave Monsanto the go-ahead for the crop trials. But such was the level of protest that the company abandoned plans to conduct the trials at three separate Irish farms, using only the Carlow research farm.

In the attack, much of the almost mature beet was destroyed, and the rest dug up. The Gaelic Earth Liberation Front, a previously unknown group, said in a statement: "This was Ireland's first genetically engineered crop and we hope it will be the last."

Patricia McKenna, a Green Party member who represents Dublin in the European Parliament, praised those responsible, adding: "If Monsanto, which was carrying out the trials, and the

Environmental Protection Agency, which licensed them, insist on playing games with the Irish environment, then fair play to those who challenge them through peaceful direct action."

Monsanto described her comment as "extraordinary, given that this was an illegal act". It said it was "shocked and dismayed at this act of wanton damage", and accused those responsible of having no interest in scientific research or the benefits it could bring. Monsanto intends to resume the trials as soon as possible.

A spokesman for Genetic Concern, an Irish environmentalist lobby which led the opposition, denied any involvement in the sabotage, or knowledge of the new group. But he added: "We're not surprised. There are a lot of people very annoyed the tests went ahead without adequate public debate."

Anthony Garvey



Shortcircuit leads India to abandon Internet satellite

[NEW DELHI] India's newest telecommunications satellite, INSAT-2D, launched only four months ago, has been abandoned following a series of problems resulting from a short-circuit in its onboard power supply. The Indian Space Research Organization (ISRO) in Bangalore said in a statement that "the satellite has now become inoperable".

The US\$50-million INSAT-2D, the fourth in a series designed and built by ISRO, was launched by Europe's Ariane rocket in June. Its loss is a major blow to India's plans to expand Internet services and provide telephone links to remote areas.

The mishap has come at a time when ISRO is already struggling to correct the orbit of a remote-sensing satellite, IRS-1D, launched two weeks ago (see *Nature* 389, 432; 1997). After an apparently perfect launch, ISRO's Polar Satellite Launch Vehicle placed the satellite in an elliptical path, bringing the spacecraft to within 300 km of Earth, instead of placing it in the circular orbit necessary for remote sensing.

ISRO has managed to push the perigee to 400 km, and says that it is confident that it can increase it further to a circular orbit from which it will be able to carry out most of its remote-sensing functions. **K.S. Jayaraman**

Asia-Pacific forum backs life science network

[TOKYO] A move by life scientists in the Asia-Pacific region to form a network to promote biotechnology and biomedical research has won strong political endorsement from the Asia-Pacific Economic Cooperation (APEC) forum (see *Nature* 388, 3; 1997).

At a meeting of APEC's working group on industrial science and technology in Singapore last week, a motion by South Korea to support the network won backing from eight APEC members: Japan, Australia, Singapore, Malaysia, the Philippines, Indonesia, Thailand and Taiwan.

Gurinder Shahi of the International Vaccine Institute in Seoul, who attended the meeting as a representative of the secretariat of the proposed International Molecular Biology Network for Asia and the Pacific Rim, says delegates from the United States and New Zealand also informally expressed support.

APEC's endorsement will not automatically be translated into financial backing, as the forum operates on a very limited budget. But it will give the network the political authority to begin work on a draft intergovernmental agreement, modelled along the lines of that setting up the European Molec-

ular Biology Organization (EMBO), many of whose characteristics the new network hopes to emulate.

Even with APEC's support, the experience of the vaccine institute suggests that it could take a long time to establish and ratify such an agreement. Meanwhile, funds are being sought from other sources — perhaps some of the institutes in the region that have expressed their support, as well as the private sector.

A task force of seven scientists from Japan, South Korea, Taiwan, Singapore, China, Hong Kong and New Zealand will meet next month in Shanghai, at the Institute of Biochemistry of the Chinese Academy of Sciences, to discuss fund-raising and future strategy.

The network's founding members favour membership based on individuals selected for their excellence, rather than on institutions, as initially proposed at a meeting in Tokyo in June. Each country will probably propose about 10 members, and EMBO may be called in to advise on the final membership list. Shahi says it is hoped that the membership issue will be settled before a conference is held by the network in South Korea next June. **David Swinbanks**

Head of Third World biotechnology centre in Delhi resigns

[NEW DELHI] The International Centre for Genetic Engineering and Biotechnology (ICGEB) is looking for a new head for its New Delhi component, following the resignation of Krishna Kumar Tewari from the post he has occupied since the unit was set up a decade ago.

Although his term expires next June, Tewari says he has resigned now to give the centre more time to choose his successor. But another factor appears to have been the tension between Tewari and the Indian government over his teaching commitments at the University of California at Irvine.

Tewari's departure was announced during a meeting last week of ICGEB's council of scientific advisers, held to review the centre's research activities in New Delhi and its twin component in Trieste, Italy.

Despite Tewari's surprise departure, ICGEB officials were upbeat last week. "The centre, which was created amidst much scepticism and mistrust, is now a centre of excellence," says director Arturo Falaschi. Forty-one countries are full members of ICGEB, and 200 scientists from 31 countries work in its two laboratories.

The centre plans to celebrate its tenth

anniversary with a symposium in Trieste in November. "ICGEB has now reached a stage where it can demonstrate that basic research carried out at its two laboratories can be exploited," says Xiaocheng Gu of the College of Life Sciences of Peking University, Beijing, who chaired last week's meeting.

Products awaiting field trials include an antimalaria peptide vaccine and an insect-resistant rice plant. A technique for producing cheap recombinant hepatitis-B vaccine will soon transfer to industry.

ICGEB trains 400 scientists from developing countries each year. And ICGEBnet, a central biocomputing resource in Trieste, provides links with 874 scientists in member countries via the Internet.

According to Falaschi, funding has improved to a point where "we are sure of maintaining the current level of activities for at least five years". But he says one weak point is the inability to obtain support from industrialized countries. Another worry is the substantial funds that will be needed to replace equipment and to expand laboratory space over the next two or three years.

Tewari had been chairman of the molecular biology and biochemistry department



ICGEB in New Delhi (above) has a twin centre in Trieste, and seeks to expand its funding base.

at Irvine for 15 years when he was appointed to head ICGEB's Indian unit. He says he will return to Irvine.

Although Tewari's direct approach to management has long irked government bureaucrats, his main problems began in May when the Department of Biotechnology objected to his spending four to six months a year teaching at Irvine. Tewari denies that his laboratory suffered from his absence, and says that ICGEB has benefited from his international connections. **K. S. J.**

Plant genome initiative wins \$40m support from Congress

[WASHINGTON] The US Congress has agreed funding for a new \$40 million-a-year plant genome sequencing effort to be led by the National Science Foundation (NSF). The initiative was proposed by Senator Christopher Bond (Republican, Missouri) and has strong backing from the agricultural biotechnology industry (see *Nature* 388, 313; 1997). It was fully funded in an appropriations bill that was agreed last week between the House of Representatives and the Senate.

The bill instructs the agency to spend \$40 million on the initiative in the new financial year, which started on 1 October, in addition to the \$20 million it had already planned to spend on plant genome research. The bill, which also increases the NSF's total expenditure on research grants by 5 per cent, or \$114 million, to more than \$2.5 billion, is expected to pass into law later this month.

Paul quits AIDS office to return to the lab

[WASHINGTON] William Paul, director of the Office of AIDS Research (OAR) at the US National Institutes of Health (NIH) since 1994, announced last week that he would step down as early as next month, having seen put in place many measures that he felt "passionately" were needed. "I leave the OAR with a true sense of accomplishment," Paul said. But, he added, "I really felt the need to return to full-time lab work. I want myself to be very much involved in the vaccine effort." Paul, who is 61, said that he plans to pursue AIDS vaccine research in his laboratory at the National Institute of Allergy and Infectious Diseases.

Harold Varmus, the NIH director, said Paul had provided "exemplary scientific leadership for the NIH AIDS research programs". Under Paul, the number of AIDS research grants grew by 50 per cent. Plans were laid for a new NIH AIDS vaccine research centre, construction of which will begin next year. But Paul's tenure was dogged by attempts by conservative Republicans in Congress to emasculate the OAR by channelling NIH's \$1.5-billion AIDS research budget directly to NIH institutes, rather than through the OAR.

Russian researchers mount funding protest

[MOSCOW] About 250 scientists picketed Russia's House of Government last week demanding that funding for science be increased from 2.85 per cent to 4 per cent of total government expenditure, as required

under current legislation. They were also protesting against plans to cut by two-thirds the number of employees in the Russian Academy of Sciences. "Ignoring our demands will destroy the nation's science and, as a result, its intellectual potential," said Vladimir Khlebodarov, chairman of the academy workers' trade union. Protest action was also held or is planned at many Russian scientific centres.

US academy president goes for second term

[WASHINGTON] Bruce Alberts, the president of the US National Academy of Sciences, has decided to put his name forward for a second six-year term. Alberts, who wants to put sustainable development and school education at the top of the academy's agenda (see *Nature* 388, 819; 1997), told its ruling council late last month that he would like to stand again.

The council will now form a nominating committee to decide which names to put before the academy's 1,800 members for the presidential election, which will take place late next year. But precedent suggests that there will only be one name — and that it will be that of Alberts.

CNRS appointments 'put science first'

[PARIS] Sixty per cent of the budget and personnel of France's National Centre for Scientific Research (CNRS) has come under new management after three appointments announced last week by Catherine Brechignac, the new director-general of the largest research organization in Europe.

The new head of life sciences is Jacques Samarut; Marie-Claude Maurel is to take charge of humanities and the social sciences; and the successor to Brechignac herself as head of physical and mathematical sciences is Jean-Paul Pouget. Brechignac says that all three are scientists appointed with the goal of ensuring that science — rather than management or politics — dominates the agenda of their departments. Such an objective, she says, reflects her priorities for CNRS as a whole.

Monster price paid for *Tyrannosaurus* Sue

[LONDON] 'Sue', claimed to be the most complete specimen known of *Tyrannosaurus rex*, fetched US\$7.6 million at Sotheby's in New York last week. The money was paid by a consortium made up of the Field Museum of Natural History in Chicago, McDonald's Ronald McDonald House Charities, Walt Disney World Resorts and others. The museum will put Sue on display from 2000, but the public will be able to see her during

preparation, says Peter Crane, the director of the museum.

Sue was discovered in 1990 on the Cheyenne River Reservation in South Dakota. The Black Hills Institute for Geological Research is said to have paid Maurice Williams, a Sioux Indian who is a ranch owner, US\$5,000 to excavate the fossil. But Sue was confiscated by the federal government, which claimed that as Williams held the land in trust he did not have the right to sell fossils found on it. The issue drew attention to the activities of commercial collectors on public land; Peter Larson, president of the institute, was jailed for two years for offences connected with the fossil trade. He is now free.

Next director lined up for Los Alamos lab

[WASHINGTON] The next director of the Los Alamos National Laboratory in New Mexico will be John Browne, a nuclear physicist who is at present director of the Los Alamos Neutron Science Center, a research facility used by both weapons scientists and outside university researchers.

The regents of the University of California were expected to confirm Browne as their choice for the position earlier this week. Browne spent 10 years at the Lawrence Livermore National Laboratory in California before arriving at Los Alamos in 1979.

Embryonic inspiration for fashion collection



[LONDON] A fashion collection inspired by the first 1,000 hours of the development of an embryo is being previewed this week at the Institute of Contemporary Arts (ICA) in London before its catwalk debut later this year. 'Primitive Streak' is the work of designer Helen Storey and her sister Kate, who is a developmental biologist at the University of Oxford. Each of the 35 outfits represents a specific stage in the development of the embryo.

The picture above shows a red and black DNA dress with a silver-plated female spine. It correlates to bone formation at about day 42. The collection grew out of a proposal funded by the Wellcome Trust under its Science and Art initiative.

Monkey business in space

Sir—The News item on NASA's Bion programme of flying restrained and heavily instrumented rhesus macaques into space portrayed it in a light that the programme does not deserve (*Nature* 387, 4; 1997). In fact, Bion's design is such that meaningful results about "how musculoskeletal and regulatory mechanisms respond to and recover from space flight" (Bion 11/12 Integrated Science Proposal and Discipline Proposals, 1996) are virtually precluded.

Both restraint and microgravity result in osteopenia and disuse atrophy of muscles, which makes it impossible to determine the specific impact of microgravity (or any other cosmic factor) on bones and muscles in the restrained animals sent into space.

The 'regulatory physiology' protocol contains no hypotheses, which, in conjunction with the interference from restraint and psychological stress, the sample size of two, and dramatic individual differences in response (as documented by previous missions), reduces Bion to 'parametric tinkering', generating a spiral of

inconclusive animal studies that are only vaguely, if at all, applicable to free-moving, relatively comfortable astronauts.

To be most useful for future missions, experimental data of this nature should be obtained from humans. The head-down tilt bed rest model simulates the physiological effects of microgravity and accurate non-invasive techniques, and osteoblast cultures are now available to study the cellular and biochemical mechanisms of osteopenia. All the musculoskeletal tests and eight out of nine 'regulatory physiology' tests could have been done in consenting humans. Eight measurements taken from astronauts are far more valuable than nine from strait-jacketed monkeys in pain and distress.

We believe that sending animals into space is one of the most ineffective ways of studying the effects of microgravity on astronauts, and that subjecting conscious mammals to extreme trauma in the name of technological progress is unethical.

Bion involved some of the most invasive instrumentation ever imposed on a

conscious primate. In addition to extensive wiring of the body and 10 incisions (including a perforation of abdominal muscles), each of the two flight and 10 control macaques had 11 holes (eight for Evarts crown, three for epidural cannulae) drilled in the skull roof and two depressions and canals in the zygomatic bones for an electro-oculogram. As well as the pain and terror of being strait-jacketed in these circumstances, the flight animals were exposed to motion sickness and shock from the dramatic concomitance of launch and landing. After returning to Earth, one macaque choked to death under anaesthesia.

David O. Wiebers

Andrzej Elzanowski

Paul W. Gikas

Jennifer Leaning

Roger D. White

Scientific Advisory Council to

The Humane Society of the United States,

2100 L Street NW,

Washington, DC 20037, USA

e-mail: elzanow@ix.netcom.com

Bet you didn't know that

Sir—Contrary to Richard Wassersug's declaration in News and Views (*Nature* 388, 827–828; 1997), the penis of the nine-banded armadillo is not the first biological structure known to have an investment of orthogonal fibres instead of the helical arrangement classical in a hydrostat.

The sac of dura mater containing the human cerebrospinal fluid and spinal cord has an orthogonal array of elastin fibres that was demonstrated by electron microscopy eight years ago (B. R. Fink & S. Walker, *Anesth. Analg.* 69, 768–772; 1989).

In the spinal dura the orthogonal pattern apparently helps to accommodate the increase in longitudinal tension occasioned by flexion of the vertebral column. The orthogonal investment of the penis, however, is composed of collagen fibres which, as Wassersug pointed out, provide the penis with maximum flexural stiffness when erect.

Raymond Fink

*Department of Anesthesiology,
University of Washington,
Seattle, Washington 98195, USA*

Definition of a gene

Sir—The term 'gene' is widely used by biologists and is also in the public vocabulary, yet it has no precise universally accepted molecular definition.

Some authors consider that a gene includes the regulatory sequences required for its expression¹. But genes do not have to be expressed to be present. The somatic cells of a multicellular organism all have the same genes, but particular cell types express only some of them.

Inclusion of regulatory sequences expands the term 'gene' from a specification of 'what it is' to indicate also 'how it is used'. It may be useful to separate these two concepts. It is thought that major evolutionary changes in the morphology of organisms have resulted from altered patterns of synthesis of the same gene products². It seems reasonable to consider that the genes themselves have not changed, but rather the ways in which they have been used.

Inclusion of regulatory sequences introduces considerable complexity to the term 'gene'. There are many different types of regulatory elements, and they generally operate in complex combinations. Some of them, such as enhancers, are nonspecific and influence any compatible promoter within their range.

Others, such as the promoters of polycistronic mRNAs, control the synthesis of several gene products. Furthermore, there is already an acceptable term ('operon') for a unit of gene expression (page 339 of ref. 3).

Should introns be considered as parts of genes? When first discovered, genes with introns were described as 'interrupted'. This

implies that a gene consists only of its exons; otherwise it would be continuous. An 'exons only' definition highlights the evolutionary theory that new genes can be created by shuffling existing exons into new combinations. There is already an acceptable term ('transcription unit') for the entire region, including introns, between a promoter and a transcriptional terminator (page 366 of ref. 4). In cases of alternative splicing of primary transcripts, it has been suggested that, unless the protein products are very different, one gene can encode a series of protein isoforms (page 457 of ref. 4).

In my opinion, the single best molecular definition of the term 'gene' is the following: it is the nucleotide sequence that stores the information which specifies the order of the monomers in a final functional polypeptide or RNA molecule, or set of closely related isoforms. This definition is simple and concise. Geneticists can readily find other names for more complex genetic entities such as 'operons' and 'transcription units'.

Christopher D. Epp

*College of Sciences,
Massey University,
Palmerston North,
New Zealand*

1. Lodish, H. *et al.* *J. Mol. Cell Biol.*, 3rd edn, 308–310 (Scientific American Books, 1995).

2. Pennisi, E. & Roush, W. *Science* 277, 34–37 (1997).

3. Lewin, B. *Genes VI* (Oxford Univ. Press, 1997).

4. Alberts, B. *et al.* *Molecular Biology of the Cell*, 3rd edn (Garland Publishing, 1994).

Science studies misjudged?

Sir — Gottfried and Wilson¹ remark critically that “SSK [sociology of scientific knowledge] accounts often treat only the earliest phases of a scientific development, when the evidence is uncertain, and largely ignore subsequent convincing confirmations”. This is true of controversy studies in SSK, but not of one of the books at issue, my own *Constructing Quarks*². There I documented and analysed the history of the establishment of the standard model as the ‘new orthodoxy’ in elementary particle physics.

Ellis³ makes a similar mistake, and adds that he would “also like [me] to document and compare more thoroughly the rejection by the scientific community of false ‘discoveries’”. I have published essays on discredited ‘discoveries’ of magnetic monopoles and isolated quarks, and on the ultimately unsuccessful theoretical arguments that the new particles (the J/ψ and so on) were manifestations of quark colour rather than charm (see, for example, ref. 4). The analysis in those essays goes along the same lines as that of *Constructing Quarks*, and if Ellis wants to dispute it, I will be happy to respond.

Capasso⁵ is mistaken in asserting that successful technology “validates the specific theories on which it is founded”. A machine works because it works, not because of what anybody thinks about it⁶. And, as I believe Gottfried and Wilson recognize, Capasso’s idea that technology owes everything to prior science is as historically misleading as its inverse. In short, the continuing criticism of science studies in your pages remains wide of its mark.

Andrew Pickering

Department of Sociology,
University of Illinois
at Urbana-Champaign,
Urbana, IL 61801, USA
e-mail: pickerin@uiuc.edu

1. Gottfried, K. & Wilson, K. *Nature* **387**, 545–546 (1997).
2. Pickering, A. *Constructing Quarks: A Sociological History of Particle Physics* (Univ. Chicago Press, 1984).
3. Ellis, J. *Nature* **388**, 13 (1997).
4. Pickering, A. *Isis* **72**, 216–236 (1981).
5. Capasso, F. *Nature* **387**, 544–545 (1997).
6. Pickering, A. *The Mangle of Practice: Time, Agency, and Science* (Chicago Univ. Press, 1995).

Narrowness in science

Sir — I agree wholeheartedly with Mott T. Greene’s assessment of narrowness in science (*Nature* **388**, 619–620; 1997). However, the narrowness goes much deeper than Greene implies. For example, my own

research area stretches from condensed-matter physics to physical chemistry, both theory and experiment. These interests are very broad by modern standards, but would probably still be classed as narrow by Greene.

Some of this narrowness has been brought on by the twentieth-century growth of science itself. In the nineteenth century, so little was known that scientific minds could comfortably stretch from medicine to astronomy. However, in addition to the scientific growth during the past half-century, science funding policies in the United States and elsewhere have provoked narrowness through the absurd and self-destructive ‘anonymous peer-review’ system. So-called peer reviewers feel that outsiders cannot know enough to contribute to their field. They thus tend to vote down financial support of any scientist who would like to broaden out, and thus encroach on the reviewer’s territory.

These trends have resulted in small knots of hostile specialists in ever-narrowing research areas. When a physical chemist is asked what field he or she is in, the usual answer nowadays is “infrared spectroscopy” or “NMR” or “statistical mechanics” or “quantum mechanics”, when the answer should be “Right at the moment, I am working on such and such a problem using a variety of these research tools”.

It is like asking a carpenter what his field is and getting the reply, "hammer" or "saw" or "screwdriver". But what do you build? What do you create? Even if such questions did occur to scientists, the problem to be solved would require such a wide variety of techniques and encroach on so many areas of specialization that, through the peer-review system, it could never be funded.

This means unfortunately that anyone with Greene's "generalized curiosity" is now unable to prosper in science. In the near future, those who go in and stay in will mainly be those with such limited thinking ability that nothing scientifically really new will ever be discovered. Is that what everyone wants?

G. Wilse Robinson

Department of Chemistry and Biochemistry,
Texas Tech University,
Lubbock, Texas 79409-1061, USA
e-mail: gwrob@ttacs.ttu.edu

Sir—I disagree with Greene's main conclusions. In my opinion, the main impediment in discussing areas of research other than our own is the knowledge that we do not know enough. While I am giving classes about CD4-gp120 interactions and CCR5/CXCR4 discrimination of HIV tropism, new receptors are being described and my knowledge seems completely outdated.

The 'information network' is so immense that we do not have time to read all about our narrow specialties, much less related areas, and still do our own research. And the pressure to publish is still increasing, the number of reviews published appears to me to be decreasing, and my ignorance of results from other areas of research stupefying.

Working in research I do not like being 'outdated' or simply wrong, so I avoid public statement of opinions.

Vera Bongertz

Lab AIDS & Molecular Immunology,
Oswaldo Cruz Foundation,
Rio de Janeiro, RJ, Brazil
e-mail: bongertz@gene.dbbm.fiocruz.br

Duplications in nomenclature

Sir—The nomenclature of genes and proteins in molecular and developmental biology, as discussed in a recent leading article in *Nature*¹, is often illogical and confusing. This problem is sometimes compounded when two proteins without any structural and functional relationship receive the same designation.

In a recent publication by Pan *et al.*², for example, the authors describe the cloning and characterization of a new

membrane-anchored chemokine and propose the name "neurotactin" for this type of molecule.

The term "neurotactin" was, however, used previously to describe a *Drosophila* membrane protein with an extracellular serine esterase-type protein domain which is dynamically expressed by neuroblasts and other tissues in the fly embryo^{3,4}. *Drosophila* neurotactin has no sequence or functional similarity to the molecule characterized by Pan *et al.*².

As a solution to such problems, journals should require the authors of manuscripts in which new names or terms are proposed to carry out a computer literature search.

In the case described above, a Medline search would have revealed the duplication and avoided possible confusion to some readers. Luckily, the CX3C membrane-bound chemokine described by Pan *et al.*² is identical to a molecule for which Bazan *et al.*⁵ proposed the name "fractalkine", so an alternative designation is available.

Michael Hortsch

Department of Anatomy and Cell Biology,
University of Michigan,
Ann Arbor, Michigan 48109-0616, USA
e-mail: hortsch@umich.edu

1. *Nature* **389**, 1 (1997).

2. Pan, Y. *et al. Nature* **387**, 611–617 (1997).

3. Hortsch, M. *et al. Development* **110**, 1327–1340 (1990).

4. De La Escalera, S. *et al. EMBO J.* **9**, 3593–3601 (1990).

5. Bazan, J. F. *et al. Nature* **385**, 640–644 (1997).

Patterns in the sand

Paul Umbanhowar

The behaviour of vibrated granular material remains largely a mystery. A new model of the process, one which emphasizes dissipation and the random nature of grain collisions, may constitute a step forward.

In one sense the dynamics of granular material are child's play — literally so, for many of us will recall how as children we wiled away the hours sieving, pouring, mixing and piling sand at the beach or in a sandbox. But how difficult is it to describe the behaviour of this common and commonly used material? The answer — as any geologist, chemical engineer or other researcher concerned with granular media will tell you — is very difficult. In an attempt to make headway, current research centres on the simplest of systems, a shaken container of sand being a popular choice¹. Yet, even here, unexpected and complex wave patterns of squares, stripes and hexagons emerge².

On page 574 of this issue³, Troy Shinbrot details a new model of vibrated granular material that reproduces these beautiful patterns. More significantly, it provides fresh insights into the fundamental physical mechanisms responsible for collective motion in granular media.

Granular wave patterns form in a thin layer of non-cohesive beads ('sand') placed on the bottom of a rigid, vertically oscillated, evacuated container². The sinusoidal shaking is characterized by the vibration frequency and peak container acceleration. Granular wave patterns arise when the peak acceleration is greater than about 2.5g. The patterns

at onset are squares for low shaking frequencies but change to stripes as the shaking frequency is increased (see Fig. 1). Hexagons appear if additional shaking at half the fundamental frequency is introduced, either externally or internally by the layer/container dynamics.

In his new model, Shinbrot discards some seemingly important items: the container, gravity, the vertical dimension and friction. What's left is a bunch of inelastic disks (32,767 to be exact) sliding about on the surface of a torus. Rather than banging into one another as real grains do, the disks interpenetrate. Collisions are then treated in an average sense by considering all the disks that overlap a 'colliding' disk as a single particle with momentum equal to the group average. These effective collisions are still inelastic, so, to keep things moving, all the disks are given simultaneous periodic kicks in random directions. In fact, the magnitude and frequency of these kicks determine the behaviour. To compare simulation with experiment, local disk density is equated with layer amplitude.

Surprisingly, this bare-bones approach recreates many of the patterns seen in experiment, including squares, stripes and hexagons. As Shinbrot explains³, the mechanism behind the formation of patterns in the simulation is the dissipation-induced clustering of

particles colliding along lines of symmetry.

For example, imagine arranging the model disks two-to-five across in a randomly but tightly packed straight line, and that several such lines run parallel to each other. A single application of random kicks leads to the effective diffusion of particles away from the original lines and into the low-density region between lines. Here counter-propagating disks from adjacent lines collide and cluster, which creates a new pattern displaced by half a wavelength from the original. (It is an interesting exercise to construct other disk distributions that lead to ordered patterns under this transformation.)

The model also exhibits a transition from squares to stripes. A similar transition is seen in the experiment when the period is decreased at constant peak acceleration. Shinbrot shows that, in the model, the square-to-stripe transition is defined by a constant ratio of the dissipated power to the input power: squares appear when the ratio is less than one, whereas stripes are seen when the ratio is greater than one.

Shinbrot's new model stands out because little is known about what causes patterns to form in granular media. In other driven dissipative nonlinear systems, such as hydrodynamic flows and chemical systems, the basics of pattern formation are well understood⁴. To determine when and why patterns arise, the stability of a base state is analysed by examining the partial differential equations that describe the system. In a vibrated granular layer, the base state is a flat layer. However, its stability cannot be calculated because there are no equations that accurately describe flowing granular material¹.

Why do such equations not exist? The difficulty is not that granular materials are composed of discrete objects — so ultimately are fluid and chemical systems. Rather, the requirements to implement a continuum

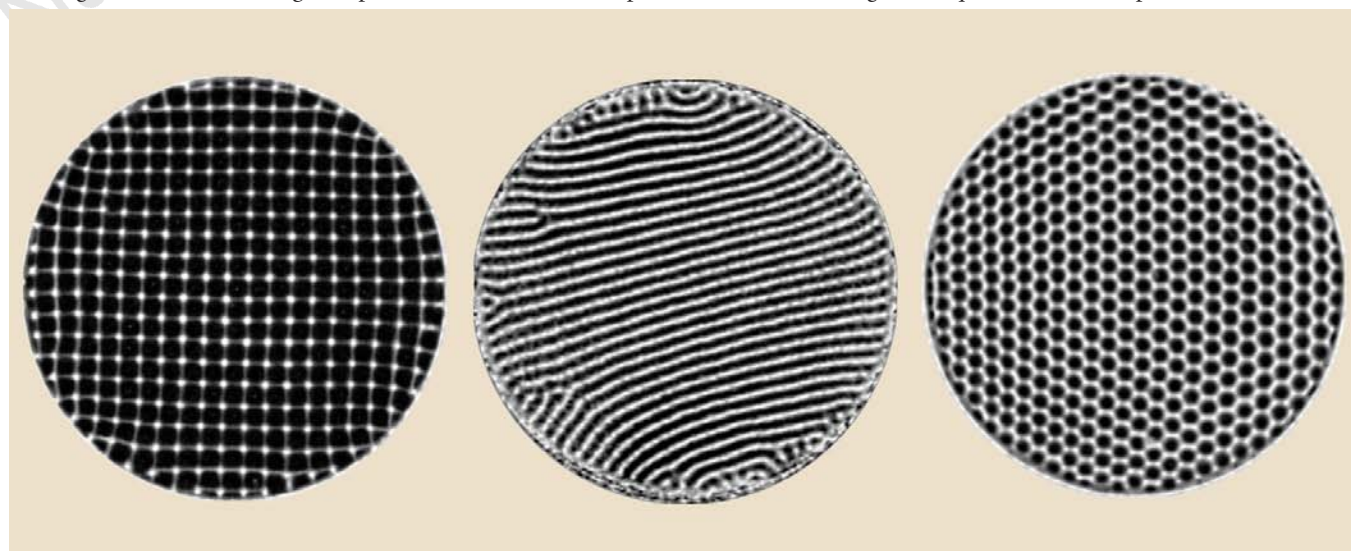


Figure 1 Square, striped and hexagonal patterns in a shaken 0.1-cm-deep layer of 0.02-cm-diameter bronze beads. The patterns are created at respective frequencies and acceleration amplitudes of (25 Hz, 3.0g), (35 Hz, 3.5g) and (32 and 16 Hz, 2.8g); container diameter is 13 cm. Compare these patterns with those in Fig. 2 (page 575) produced in the model simulations³ discussed here.

description are lacking. First, the particle size is large compared with the length scale over which the flow varies (on the order of tens of particle diameters for granular waves). Second, thermal energy is too small to move particles around, and thus an undriven system is effectively 'frozen' in a non-unique metastable state. Giving the grains a kick does not solve this problem, because the subsequent collisions dissipate the added energy in less time than the time needed to equilibrate the system spatially.

Despite these obstacles to developing descriptions of granular flow through conventional methods, an understanding of vibrated granular patterns has been emerging thanks to two different approaches. The first is the microscopic, particle-level simulation of two-dimensional⁵⁻⁷ and three-dimensional⁸ systems. These simulations reproduce granular patterns, and they show that existing grain-interaction models are sufficient to describe vibrated layers. When performed accurately, simulations also allow the measurement of quantities that are difficult to obtain experimentally. There remains, however, the task of distilling general properties of granular dynamics from the simulations.

The second approach uses continuum phenomenological descriptions to characterize vibrated granular systems⁹⁻¹¹. These models capture important pattern behaviour, such as the square-to-stripe transition⁹ and even the existence of recently discovered 'oscillons' (highly localized standing waves)¹². Unfortunately, it is not clear how to determine the parameters which characterize these models from measurable grain properties; direct connections with experiment are therefore difficult to make.

Shinbrot's model is essentially a hybrid of these two approaches. It suggests what physical mechanisms in the particle-level simulations are relevant, and, through its use of mean-field-like particle interactions, points to possible analytical approaches to connect phenomenological continuum models to the discrete system. However, although the model reproduces many of the patterns and qualitative behaviours found in vibrated sand, as Shinbrot acknowledges, it resembles a chemical reaction-diffusion system more than a granular flow. Both physical systems produce similar patterns.

An exciting opportunity to test whether or not the model truly captures the essence of granular pattern formation is provided by the prediction of a new pattern not yet seen in experiments. The pattern is composed of isosceles triangles arranged tip to tail in alternating rows (see upper right image in Fig. 2 on page 575), and appears in the simulations at low driving frequencies and large kick velocities. This combination of parameters is challenging to access experimentally because large shaking amplitudes are necessary to

generate high velocities at low frequency. But a laboratory observation would certainly provide convincing evidence that pattern formation in vibrated granular materials depends primarily on periodic kicks and the random, dissipative nature of grain collisions. □

Paul Umbanhowar is in the Physics Department, University of Pennsylvania, David Rittenhouse Laboratory, 209 South 33rd Street, Philadelphia, Pennsylvania 19104, USA.

e-mail: pbu@dept.physics.upenn.edu

1. Jaeger, H. M., Nagel, S. R. & Behringer, R. P. *Physics Today* **49** (11), 32–38 (1996).

Kainate receptors

Finding homes at synapses

Mark Mayer

Synapses are the structures that the brain uses to process information and store memories. Textbook descriptions of a typical excitatory synapse show glutamate — released from presynaptic nerve terminals — activating two families of receptors, the AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) and NMDA (*N*-methyl-D-aspartate) receptors, in the postsynaptic cell. The function of a third family of glutamate receptors, which have a high affinity for the neurotoxins kainate and domoate, has remained an enigma. Anatomical studies show that kainate-receptor messenger RNA and protein are widely expressed in brain tissue, yet attempts to identify synaptic responses linked to kainate receptors have been unsuccessful. But now, a series of papers¹⁻³ culminating in the letter by Clarke *et al.*¹ on page 599 of this issue, reveal, for the first time, a role for kainate receptors in synaptic transmission.

Why did these new studies succeed in finding kainate-receptor responses where previous work has failed? The answers are, in retrospect, both predictable and surprising. New pharmacological probes were pivotal in allowing Clarke *et al.* to identify a presynaptic inhibitory effect by kainate on the release of another neurotransmitter, GABA (γ -aminobutyric acid). So, with better tools in the pipeline, it should soon be possible to identify additional functions for kainate receptors at other synapses. Perhaps more surprising was the identification by Castillo *et al.*², and Vignes and Collingridge³, of kainate-receptor excitatory postsynaptic responses. This was achieved by a combination of persistence and clever experimental design, but using tools that had been available for several years.

To understand better why this work is of interest, we need to step back almost 20 years to when kainate receptors were first identified as a distinct class of receptors⁴. Support for this view came from observations⁵ of the uniquely high density of kainate-binding

2. Melo, F., Umbanhowar, P. B. & Swinney, H. L. *Phys. Rev. Lett.* **75**, 3838–3841 (1995).
3. Shinbrot, T. *Nature* **389**, 574–576 (1997).
4. Cross, M. C. & Hohenberg, P. C. *Rev. Mod. Phys.* **65**, 851–1112 (1993).
5. Luding, S., Clément, E., Rajchenbach, J. & Duran, J. *Europhys. Lett.* **36**, 247–252 (1996).
6. Aoki, K. M. & Akiyama, T. *Phys. Rev. Lett.* **77**, 4166–4169 (1996).
7. Rapaport, D. C. *Physica A* (in the press).
8. Bizon, C., Shattuck, M. D., Swift, J. B., McCormick, W. D. & Swinney, H. L. *Phys. Rev. Lett.* (submitted).
9. Tsimring, L. S. & Aranson, I. S. *Phys. Rev. Lett.* **79**, 213–216 (1997).
10. Cerda, E., Melo, F. & Rica, S. *Phys. Rev. Lett.* (submitted).
11. Rothman, D. H. *Phys. Rev. Lett.* (submitted).
12. Umbanhowar, P. B., Melo, F. & Swinney, H. L. *Nature* **382**, 793–796 (1996).

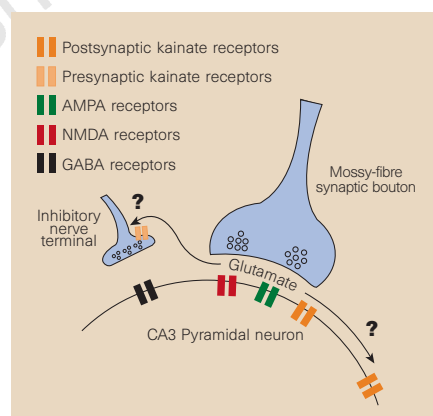


Figure 1 A new biology for kainate receptors at central synapses. Postsynaptic sites on CA3 pyramidal neurons innervated by mossy fibres, along with presynaptic sites on GABA (γ -aminobutyric acid)-releasing inhibitory nerve terminals, seem to express functional kainate receptors. The GluR5-selective agonists used by Clarke *et al.*¹ give clues as to the subunits expressed on inhibitory nerve terminals, indicating that kainate-receptor-mediated depolarization may interfere with the activation of voltage-dependent Ca^{2+} or Na^{+} channels required for neurotransmitter release.

sites in area CA3 of the hippocampus (Fig. 1). As the cloning of glutamate-receptor complementary DNAs progressed, many gene families were identified for glutamate receptors⁶. Yet despite the successful functional characterization of cloned kainate receptors assembled from the GluR5–7, KA-1 and KA-2 subunits, synaptic responses due to activation of kainate receptors in tissue from the central nervous system could not be identified. Even attempts to show kainate-receptor responses in the brain were problematic, although functional kainate receptors were shown to be expressed in dorsal-root ganglion neurons, glial-cell O-2A progenitors, and hippocampal neurons in culture⁷⁻⁹.

Castillo *et al.*² and Vignes and Collingridge³ took advantage of the discovery

that 2,3-benzodiazepines selectively block AMPA receptors but not kainate receptors^{10,11}. By looking carefully at synaptic responses in hippocampal area CA3 — in which kainate receptors are expressed at highest density in the region where mossy fibres form synapses — both groups showed excitatory synaptic currents, which could best be explained by activation of kainate receptors. Castillo *et al.* also went to great lengths to show that these responses are produced only when mossy fibres are activated. The facilitation of synaptic responses observed at high stimulation frequencies could reflect spill-over of glutamate to adjacent synapses or, more likely, recruitment of additional release sites within the complex structure of a mossy-fibre nerve terminal. Matching these responses with individual kainate-receptor subtypes remains an important issue which may soon be addressed by the analysis of kainate-receptor knockout mice.

The discovery that activation of kainate receptors has an inhibitory effect on the release of GABA was made possible by the development of GluR5-selective kainate-receptor agonists and antagonists. Clarke *et al.*¹ show that GABA-mediated inhibitory synaptic responses in CA1 pyramidal neurons are suppressed by GluR5 agonists. Moreover, this effect is blocked by a GluR5-selective antagonist. The nonselective agonist kainate has already been shown to produce a similar depression of excitatory synaptic transmission¹².

We now need to define the mechanism by which presynaptic kainate receptors reduce neurotransmitter release. An obvious possibility (which brings to mind memories of primary-afferent depolarization — a popular model of presynaptic inhibition before G-protein-mediated inhibition of calcium channels and GABA_B receptors was discovered), would be that depolarization of nerve terminals, mediated by GluR5 receptors, inactivates the voltage-dependent Ca²⁺ channels that are required for neurotransmitter release. Another hypothesis would be that activation of GluR5 blocks the invasion of action potentials into inhibitory nerve terminals. It will also be important to establish that there are endogenous mechanisms by which glutamate can activate GluR5 to inhibit excitatory and inhibitory synaptic transmission. At present, suppression of synaptic responses by activation of kainate receptors has been shown only for experimentally applied agonists. It will be difficult to establish the source of the glutamate that is acting on the presumptive presynaptic GluR5 receptors. Using immunohistochemistry and electron microscopy, it might be possible to discover whether GluR5 is located within postsynaptic densities on GABA nerve terminals, or whether it occurs diffusely — suggesting, perhaps, activation by extracellular glutamate rather than through

a conventional synaptic mechanism.

These are exciting days for kainate receptors, and the main challenge now will be to integrate the new observations into a revised picture of how glutamate acts as a synaptic transmitter. We can also be sure that the pharmaceutical industry will be exploring the possibility that drugs which selectively block excitatory postsynaptic responses, or which, perhaps, modulate the effects of kainate on neurotransmitter release, can act as new anticonvulsants. □

Mark Mayer is in the Laboratory of Cellular and Molecular Neurophysiology, National Institute of Child Health and Human Development, NIH, Bethesda, Maryland 20892, USA.

Climate change

Volcanic sulphur in the balance

Michael R. Carroll

On page 591 of this issue¹, Victor Kress offers a new theory to account for the large sulphur emissions from the 1991 eruption of Mount Pinatubo, Philippines. The importance of this issue is that, when they reach the stratosphere, such emissions may contribute to global cooling of the atmosphere, and Kress's theory may assist in identifying other sulphur-rich volcanic eruptions that have occurred in the past. Overall, the object of this line of research is to produce realistic estimates that can help us in balancing the forces — past, present and future — of climate cooling and climate warming.

The potential influence of volcanic eruptions on the Earth's climate has long been recognized, but quantifying such effects is far from easy. Original ideas that dust from volcanoes causes cooling by blocking solar radiation have been shown to be of limited importance, because such dust stays in the atmosphere comparatively briefly². In con-

1. Clarke, V. R. *et al.* *Nature* **389**, 599–603 (1997).
2. Castillo, P. E., Malenka, R. C. & Nicoll, R. A. *Nature* **388**, 182–186 (1997).
3. Vignes, M. & Collingridge, G. L. *Nature* **388**, 179–182 (1997).
4. Watkins, J. C. & Evans, R. H. *Annu. Rev. Pharmacol. Toxicol.* **21**, 165–204 (1981).
5. Foster, A. C., Mena, E. E., Monaghan, D. T. & Cotman, C. W. *Nature* **289**, 73–75 (1981).
6. Hollmann, M. & Heinemann, S. *Annu. Rev. Neurosci.* **17**, 31–108 (1994).
7. Huettner, J. E. *Neuron* **5**, 255–266 (1990).
8. Lerma, J., Paternain, A. V., Naranjo, J. R. & Mellström, B. *Proc. Natl Acad. Sci. USA* **90**, 11688–11692 (1993).
9. Patneau, D. K., Wright, P. W., Winters, C., Mayer, M. L. & Gallo, V. *Neuron* **12**, 357–371 (1994).
10. Paternain, A. V., Morales, M. & Lerma, J. *Neuron* **14**, 185–189 (1995).
11. Wilding, T. J. & Huettner, J. E. *Mol. Pharmacol.* **47**, 582–587 (1995).
12. Chittajallu, R. *et al.* *Nature* **379**, 78–81 (1996).

trast, volcanic sulphur emissions react to produce sulphuric acid aerosols which have much longer residence times; these aerosols can efficiently backscatter and absorb incident solar radiation, resulting in atmospheric cooling^{2–4}.

The main component of volcanic stratospheric aerosols is drops of sulphuric acid⁵, and studies of polar ice cores^{6,7} show that the aerosols are dispersed globally. The implication of these observations is that the effects of volcanic activity on climate are mainly controlled by the amount of sulphur emitted, and are not simply related to the size of the eruption or the amount of volcanic dust produced^{8,9}. Although global estimates of such emissions suggest that they are much smaller than those stemming from human activity, large explosive eruptions can inject sulphur directly into the stratosphere whereas anthropogenic sulphur is mainly limited to very low altitudes and is rapidly removed in precipitation. So volcanic sulphur could



Figure 1 Mount Pinatubo in eruption, injecting large amounts of sulphur into the stratosphere. Kress's model¹, discussed here, provides a possible explanation for the cause of sulphur-rich eruptions such as this, which may contribute significantly to global cooling of the atmosphere.

FRANK SPOONER

potentially have a much greater influence on climate, and identification of sulphur-rich eruptions in the geological record is of prime importance in trying to find links between past climatic variations and volcanic activity.

Estimating the amount of sulphur injected into the atmosphere during eruptions relies on studies of erupted materials or, in the case of recent events, on remote-sensing measurements. The latter measurements make use of satellite-based spectroscopy designed to map atmospheric ozone, and can provide estimates of the total amount of sulphur emitted during a given eruption. Studies of erupted materials involve analysis of melt inclusions, trapped within crystals present in the erupted material, which provide a measure of the pre-eruptive sulphur content of the magma — that is, the crystals act like sealed containers for samples of high-pressure melts. Comparison of the sulphur content of undegassed melt inclusions with that of the surrounding, degassed matrix glass provides a measure of how much sulphur was lost during an eruption.

Scaling this difference for the total volume of material erupted gives an estimate of the sulphur mass injected into the atmosphere. In the case of several recent eruptions involving oxidized, CaSO_4 -bearing magmas, the sulphur yields estimated by these two methods have differed by up to several orders of magnitude, the satellite-based estimates being the larger. Hence the much debated 'excess sulphur' problem in a number of recent eruptions (for example, those of Mt Pinatubo; El Chichón, Mexico; and Lascar, Chile). To try to provide an explanation, Kress's model makes use of how sulphur solubility in melts and the stability of sulphur-bearing phases — pyrrhotite (Fe_{1-x}S) and anhydrite (CaSO_4) — varies with magma oxidation state.

At Pinatubo there is evidence that the main eruption of an oxidized, anhydrite-bearing dacitic magma was preceded by mixing between a relatively reduced basalt and the predominant dacitic magma. Sulphur solubility shows a minimum at oxidation states intermediate between the states of these two mixing end-members, and the stabilities of the sulphur-rich minerals pyrrhotite and anhydrite are reduced at these intermediate states (anhydrite is not stable).

As a result, mixing of oxidized, anhydrite-bearing magma with reduced, sulphide-saturated magma may produce an intermediate magma with much lower sulphur content and liberate large amounts of gaseous sulphur (due to sulphur loss from the melt and breakdown of anhydrite and pyrrhotite). In this way, mixing could have produced abundant gaseous sulphur, and might have acted as an eruption trigger — the exsolved gas would make the mixed magma less dense than surrounding magma, resulting in gravitational instability and buoyant rise; further

decompression and gas exsolution could then precipitate an eruption.

This view of events is certainly enticing, but it is not the only possible explanation for the sulphur-rich character of the Pinatubo eruption. The underestimate of sulphur yield from studies of erupted materials could be explained by the presence of a sulphur-rich vapour phase in the magma chamber before the eruption or any mixing events. The work of Wallace and Gerlach¹⁰ suggests that this vapour phase would have comprised 6–9 per cent by volume of erupted magma. It is unclear, however, whether it is possible to concentrate and store such a large gas mass in sub-volcanic magma chambers, because only very small amounts of crystallization would be required to develop significant overpressures, thereby precipitating eruption.

The advantage of Kress's model¹ is that it can produce the required mass of sulphur comparatively quickly, and that it involves mixing of realistic masses of basaltic and dacitic magma. But a complication is that both thermodynamic and mass-balance constraints will influence the eventual outcome of mixing. For example, oxidation of a sulphide-saturated basalt will tend to consume sulphide in order to make a more sulphur-rich vapour phase; but if oxidation is due to mixing with a relatively sulphur-rich dacite, there may be enough sulphur in the dacite to provide the required sulphur-rich

vapour without the need to consume all of the basalt's sulphide. So one can imagine mixing conditions and end-member compositions that would not release large masses of gaseous sulphur.

Fortunately, such questions can be tested by more detailed petrological and geochemical studies of the erupted materials. Kress's proposal that magma mixing results in large emissions of gaseous sulphur is plausible, and further studies of eruptions involving mixed magmas should be pursued. Indeed, unpublished work by myself and colleagues on melt inclusions in mixed magmas at Lascar shows clear evidence for mixing of reduced basaltic magma (preserved in melt inclusions in xenocrysts) with anhydrite-bearing dacites — this may be yet another example of an anomalously sulphur-rich eruption, much like that of Mt Pinatubo. □

Michael R. Carroll is in the Department of Geology, Bristol University, Bristol BS8 1RJ, UK.

e-mail: mike.carroll@bris.ac.uk

1. Kress, V. *Nature* **389**, 591–593 (1997).
2. Pollack, J. B. *et al.* *J. Geophys. Res.* **81**, 1071–1083 (1976).
3. Rampino, M. R. & Self, S. *Quat. Res.* **18**, 127–143 (1982).
4. Devine, J. D., Sigurdsson, H. & Davis, A. N. *J. Geophys. Res.* **89**, 6309–6325 (1984).
5. Castleman, A. W. Jr, Munkelwitz, H. R. & Manowitz, B. *Tellus* **26**, 222–234 (1974).
6. Hammer, C. U. *J. Glaciol.* **25**, 359–372 (1980).
7. Legrand, M. R. & Dalmas, R. *J. Nature* **327**, 671–676 (1987).
8. Lamb, H. H. *Paleogeogr. Paleoclimatol. Paleocool.* **10**, 203–230 (1971).
9. Kennett, J. P. & Thunell, R. C. *Science* **196**, 1231–1234 (1977).
10. Wallace, P. J. & Gerlach, T. M. *Science* **265**, 497–499 (1994).

Linguistics

The language steamrollers

Jared M. Diamond

The world's 6,000 or so languages have a very uneven geographic distribution, in numbers both of languages and of distinct linguistic stocks per square kilometre (Table 1). At the one extreme, Europe, with an area of about 10,000,000 km², has only about 63 native languages, falling into only three stocks (51 Indo-European languages, 11 Uralic and the isolate Basque). At the opposite extreme, New Guinea, with less than one-tenth of Europe's area, has about 1,000 languages — yes, 1,000 mutually unintelligible languages, not mere dialects — falling into at least 60 stocks.

Peter Bellwood¹ and Colin Renfrew² have built on their own earlier studies^{3,4} to relate these contrasts to massive expansions of a few linguistic groups long before the Industrial Revolution or state formation, but after the end of the Pleistocene some 10,000 years ago. These early Holocene linguistic upheavals are crucial for understanding human population genetics, as well as archaeology and history^{5,6}.

Most languages fall into some family of related languages, of which Eurasia's 144

Indo-European languages and its 24 Uralic languages are examples^{7,8}. Although linguists differ over how many distinct language families they recognize for the Native Americas and New Guinea, there is widespread agreement that Africa has only four families and Eurasia about a dozen. The existence of a language family implies that its modern languages diverged from some shared ancient ancestral language, as documented for the divergence of the modern Romance languages (constituting a branch of the Indo-European family) from Latin over the past 2,000 years. By using such measurable rates of divergence to calibrate the 'time depths' of the main language families, one estimates that each of them has accumulated language divergence over about the past 3,000–10,000 years.

Remarkably, that implies that the vast majority of existing Old World (Eurasian plus African) languages are descended from a mere 16 or so languages that existed 10,000 years ago. Surely, in the early Holocene the Old World actually supported far more languages than 16 — there must have been

tens of thousands of them, if modern New Guinea or Native American California can be taken as models. Most of those ancient Old World languages must have disappeared within the past 10,000 years, leaving as evidence of their former existence only a few isolated languages that barely survived into modern times. In western Europe, only Basque and the now-extinct Etruscan and possibly Minoan (Linear A) languages attest to the linguistic diversity erased by the sweep of the Indo-European steamroller over Europe. What enabled speakers of those 16 ancestral languages to supplant their tens of thousands of brethren?

Both Bellwood and Renfrew attribute the steamrollers to the very local origins of agriculture around the world. At most only nine circumscribed areas, perhaps as few as five, supported a sufficient diversity of domesticable wild plant and animal species to permit food production to arise independently, beginning about 10,000 years ago in the Fertile Crescent of southwest Asia^{9,10}. Because even ancient food production yielded 10–100 times the human population densities that could be supported by the hunter–gatherer lifestyle, farmers spread from those few homelands to interbreed with, dominate or replace the hunter–gatherers, and thereby carry their domesticated animals, languages and genes over most other areas suitable for agriculture. By erasing the products of previous tens of thousands of years of language evolution, the Holocene agricultural expansions reset the linguistic clock in much of the world. Only a few regions, such as New Guinea and Native California, remained unaffected.

The occurrence of agricultural expansions is also supported by evidence from archaeological horizons, such as the Linear-bandkeramik (Linear pottery) horizon associated with agriculture's rapid spread over central and western Europe. Massive population replacements within the past 10,000 years are also inferred from the large biological contrasts persisting today, in some areas,

Ancdata Oxoniensia

THE
EARLIEST TRANSLATION
OF THE
OLD TESTAMENT
INTO THE BASQUE LANGUAGE (A FRAGMENT)

BY
PIERRE D'URTE OF ST. JEAN DE LUZ, *circa* 1700

EDITED, FROM A MS. IN THE LIBRARY OF SHIRBURN CASTLE, OXFORDSHIRE
By LLEWELYN THOMAS, M.A.
FELLOW OF JESUS COLLEGE, OXFORD

Hastean creatu çitu-
-en Iaincoac çerúac
eta Lúrra.

2 eta Lurra molderic
gabe çen eta hutssa
eta Íllhumbeac çiren
Leçearen gagnean
eta Iaincoaren Izpiritua
higuitçen çen vren gagne-
-an

3 eta Iaincoac erran çuen,
içan bedi argüi: eta
arguia içatu çen

4 eta Iaincoac ikhussi
çuen arguia ona çela
eta Iaincoac apartatu
çuen arguia Íllhumbe-
-enganic

...

In the beginning
God created the heaven
and the earth.

2 And the earth was
without form, and void;
and darkness was upon
the face of the deep.
And the Spirit of God
moved upon the face
of the waters.

3 And God said,
Let there be light: and
there was light.

4 And God saw the
light, that it was good:
and God divided the light
from the darkness

...

Figure 1 Basque, a living language, here in the form of a translation of Genesis. In western Europe, Basque is the sole extant survivor of the diversity erased by the Indo-European linguistic steamroller.

between descendants of evidently intrusive groups and aboriginal groups — such as the contrast between Black African farmers and Pygmy and Khoisan hunter–gatherers in sub-Saharan Africa, or between widespread Asian-like and local New-Guinean-like populations in Indonesia and the Philippines.

Each of the two earliest centres of agriculture in the Old World — southwest Asia and

China — is the inferred homeland of four or five now-widespread language families. From southwest Asia came the languages ancestral to the modern Indo-European, Dravidian, Turkic and Semitic language families or groups, while China spawned the Sino-Tibetan, Austroasiatic, Austronesian, Tai-Kadai and Miao-Yao language families. So those two homelands account for the languages spoken by about 90 per cent of all people alive on Earth today.

In thus painting the big picture, neither Bellwood nor Renfrew glosses over the many fascinating complications. The Americas included three agricultural homelands (Mexico, the eastern United States, and the Andes and Amazon). However, agriculture-driven linguistic expansions in the Americas were on a smaller scale than in the Old World, probably because of the Americas' paucity of domestic animals and the hemisphere's north–south axis. The New Guinea highlands were also an agricultural homeland and spawned at least one linguistic expansion (the Trans-New-Guinea language phylum). But New Guinea's modest range of crops and domestic animals, highly dissected terrain and isolation of highland areas by lowland

Table 1 **World variation in linguistic diversity**

Region	Area (km ²)	Languages	Language stocks	Languages per 10 ⁶ km ²	Stocks per 10 ⁶ km ²
New Guinea	800,000	~1,000	~60	1,250	75
California	411,000	~64	~6	156	15
Africa	30,000,000	1,469	4	49	0.13
Europe	10,000,000	63	3	6	0.30
Eurasia	55,000,000	808	17	15	0.30

'Language stocks' is the number of recognized language families (for instance, Indo-European) plus isolated languages not assignable to any family (for instance, Basque). All numbers refer to languages of 'aboriginal populations', the populations resident in the region concerned just before the worldwide expansion of Europeans began in AD 1492. For instance, numbers for California refer to Native American Indian tribes of California and exclude neo-American immigrants. (Data taken from ref. 7 and other sources.) Note that the density of languages or stocks per 10⁶ km² is lowest in Africa, Europe and Eurasia (because most of their former languages and stocks were submerged by agriculture-driven population expansions); and that the density of languages decreases in the sequence Africa → Eurasia → Europe (because the subsequent rise of state societies in Eurasia, especially Europe, has extinguished still more languages).

rainforest ensured the wholesale survival of early New Guinea language stocks.

Some food-producing peoples outside agricultural homelands are not the descendants of expanding farmers but of indigenous hunter-gatherers converted to food production, such as the livestock-herding Khoi population of South Africa. Conversely, some modern hunter-gatherers are related farmers, such as Great Basin Native American tribes, Polynesia's Moriori and South Island New Zealand Maori, and Austronesians of Borneo's interior. Some Holocene linguistic expansions were of hunter-gatherers at the expense of other hunter-gatherers, notably that underlying the Pama-Nyungan language family occupying most of Aboriginal Australia.

Many likely correlations between linguistics and widespread, homogeneous archaeological horizons cry out for identification. What language was spoken by the first farmers of the Fertile Crescent (Pre-Pottery Neolithic A horizon), the Linearbandkeramik invaders of Europe, the Harappan city-dwellers of the Indus Valley, the first rice farmers of India, and the Mississippian mound-building societies of the eastern United States? One might speculate: respectively, the answers could be a proto-Semitic language, a language ancestral to some Indo-European branch (but which one?), a proto-Dravidian language, a proto-Munda Austroasiatic language, and an ancestral language of who-knows-what Native American linguistic stock.

The population expansions of the Holocene have major implications for understanding human genetic history. There is great current interest in much earlier postulated population replacements, as anatomically modern humans emerged between about 150,000 and 40,000 years ago. They include the hotly debated replacements of Neanderthals and other Eurasian populations by populations derived from Africa. Geneticists approach these questions by sampling human populations on different continents. But genetic diversity within each continent has been greatly influenced by those known population expansions of the past 10,000 years, which homogenized the populations of some continents (notably, Eurasia) much more thoroughly than those of others (Africa, the Americas, New Guinea). In addressing questions about the earlier population expansions in the Pleistocene, geneticists' sampling regimes must take into account the effects of the subsequent Holocene expansions.

Suppose, for instance, that one took an African sample including populations (Pygmies and Khoisan) that had survived the Holocene expansion of Bantu farmers, and then compared that African sample with seemingly more far-flung European and East Asian population samples, all of which were

in reality recently derived from farmers' expansions out of small areas of southwest Asia and China respectively. The resulting higher calculated inter-population genetic diversity of the African sample might then tell us more about Holocene dispersals than about Pleistocene dispersals. Precisely these considerations are, of course, what underlie the proposal of a Human Genome Diversity Project — which seeks to obtain representative sampling of surviving genetic diversity before our most informative remaining populations at last become steamrolled out of their separate existence.⁶ □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90095-1751, USA.

Developmental biology

Straight and wiggly affinities

Peter A. Lawrence

During growth and development, different kinds of cells will adhere to one another; however, like oil and water, they do not mix freely. The properties that make cells more miscible with their own kind were named 'cell affinities' long ago (see ref. 1), but their molecular origins are still not well understood. Two papers, one by Blair and Ralston in *Development*², the other by Rodriguez and Basler on page 614 of this issue³, report a new and unexpected result which shows that an intercellular signal helps specify cell affinities.

These affinities become apparent when two types of cells are dissociated, mixed and allowed to sort from each other⁴, or when patterns of growth are studied *in vivo* by marking clones of cells. Clones made within populations of cells with similar affinity have wiggly boundaries, because the growing cells mingle with their neighbours. But where two populations of different cells meet, contact between the cell types is minimized and the boundary becomes relatively straight. When affinities are very different the clone may sort out *in vivo*, its perimeter becoming shorter and shorter until the clone is thrown up on a stalk that can later pinch off, leaving a separate vesicle in the body cavity⁵. Thus the boundary line between growing cells of different types, straight or ragged, is a measure of the degree of mutual affinity of the cells on either side.

This matter is central in understanding development of insect embryos, which are constructed in modules, or compartments. Where one compartment contacts another, the two populations of cells tend to form a straight interface, showing that there is an abrupt change of cell affinity at the compartment boundary. Within the compartments the clonal boundaries are wiggly. For example, the wing of *Drosophila* is made of two

1. Bellwood, P. in *The Origins and Spread of Agriculture and Pastoralism in Eurasia* (ed. Harris, D. R.) 465–498 (UCL Press, London, 1996).
2. Renfrew, C. in *The Origins and Spread of Agriculture and Pastoralism in Eurasia* (ed. Harris, D. R.) 70–92 (UCL Press, London, 1996).
3. Bellwood, P. *Sci. Am.* **265**(1), 88–93 (1991).
4. Renfrew, C. *Archaeology and Language: The Puzzle of Indo-European Origins* (Cape, London, 1987).
5. Cavalli-Sforza, L. L., Piazza, A. & Menozzi, P. *The History and Geography of Human Genes* (Princeton Univ. Press, 1994).
6. Cavalli-Sforza, L. L. & Cavalli-Sforza, F. *The Great Human Diasporas* (Addison-Wesley, Reading, MA, 1995).
7. Ruhlen, M. *A Guide to the World's Languages* (Stanford Univ. Press, 1987).
8. Comrie, B., Matthews, S. & Polinsky, M. *The Atlas of Languages* (Facts on File, New York, 1996).
9. Diamond, J. *Guns, Germs, and Steel* (Norton, New York, 1997).
10. Smith, B. D. *The Emergence of Agriculture* (Sci. Am. Libr., New York, 1995).

compartments, the distinctions between the two depending on the selector gene *engrailed* which is active in the posterior (P) compartment and inactive in the anterior (A) one⁶.

The hypothesis that the different affinities of A and P cells depend on the *engrailed* gene being responsible for 'labelling' of P cells was published in 1975 (ref. 6). Some

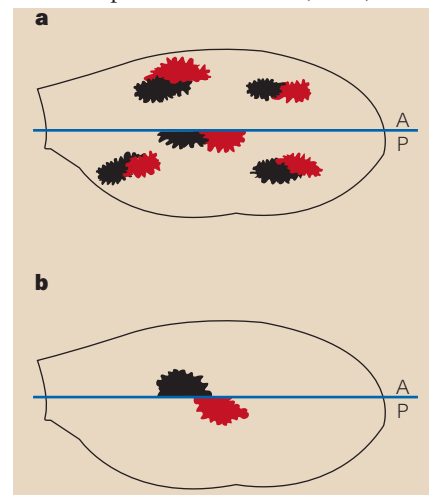


Figure 1 Analysis of cell affinities in the developing wing of *Drosophila*, which is divided into anterior (A) and posterior (P) compartments. The experiments^{2,3} discussed here involve the generation of pairs of marked clones, or 'twin spots', with (*smo*⁺, black) and without (*smo*⁻, red) the *smoothed* gene. a, The usual situation, in which clones in the two compartments never cross the A/P divide; note that the *smo*⁻ sisters in A are far from the boundary and stay where they are. b, Both groups^{2,3} find that when a *smo*⁻ sister, which has A identity, is made near to the A/P border, it moves into P. Their observations differ however as to the nature of the boundaries the clone then has with its neighbours.

experiments support this view: for example, if the *engrailed* gene is removed from the P cells, the cells make structures characteristic of the A compartment (that is, they acquire A identity); they also gain different affinities and they sort out from their neighbours, even forming vesicles in some cases. If such cells are made near to the A/P border they can sort into the A compartment and define the border from the anterior side. So there is little doubt that *engrailed* controls the differences between A and P cells, including their affinities.

The two groups^{2,3} now delve into the question of whether *engrailed* specifies affinities directly or indirectly. In the growing imaginal wing disk, P cells signal by means of Hedgehog protein (Hh), which instructs the nearby A cells to turn on synthesis of Decapentaplegic (Dpp), a long-range morphogen; Hh also gives the cells positional information by telling them that they are near the border (see ref. 7). The experimental method in both papers is the same, and makes use of the *smoothened* gene; this gene encodes a component of the Hh receptor which the cells depend upon to see Hh (ref. 8). Flies are made in which the sister cells from a single division can be individually marked to generate a pair of marked clones, a 'twin spot'. One sister clone lacks the *smoothened* gene (*smo*⁻), and the other (*smo*⁺) is wild type (Fig. 1a).

Both papers include beautiful pictures showing that when a twin spot is made in the A compartment, near the boundary, the *smo*⁻ clone, which retains A identity, nevertheless moves backwards into P territory, while its sister *smo*⁺ clone stays in the A domain (Fig. 1b). Moreover the *smo*⁻ clone now forms a straight edge against the other A cells, close to the normal position of the A/P boundary. This is a remarkable finding; it suggests that the largest contributor to the affinity difference between A and P cells, at least in the region of the A/P boundary, is not Engrailed acting directly on the P cells, but is an indirect result of Hh signalling from P to A cells across the border.

There is, however, a point of difference between the two sets of results. Rodriguez and Basler show the *smo*⁻ clones forming a reasonably wiggly boundary with their new P neighbours in the imaginal disk, whereas Blair and Ralston look at the pupal wing and find that the displaced *smo*⁻ clone has rather straight boundaries with both A and P neighbours. These results then take the authors in different directions: Rodriguez and Basler argue that Engrailed acts indirectly on affinities entirely, or almost entirely, through the Hh signal; Blair and Ralston suggest that, in addition, Engrailed may act autonomously to give a distinct affinity to P cells.

To this mix I would add one point. There is some evidence that affinities are graded along the antero-posterior axis⁹, making the

positional identity of the *smo*⁻ cells crucial to the argument. One would expect that the mutant cells, not receiving any Hh, would 'believe' they were at the anterior edge of the A compartment — that is, remote from the Hh source *in situ* and, therefore, also remote from the peak of the Dpp gradient. In the adult abdomen of *Drosophila*, *smo*⁻ clones located in an equivalent position do indeed make the type of cuticle found remote from the A/P boundary (ref. 10, Figs 6B, 7).

If the Hh signal were the only effector of affinity, the *smo*⁻ cells, receiving no signal, should retain the same affinity as the P cells, and form a wiggly boundary with them, and this is what Rodriguez and Basler show in the disk. By contrast, Blair and Ralston's results show that the *smo*⁻ cells form straight boundaries with both A and P neighbours in the pupal wing, and this also happens in the abdomen. These observations seem to support each other, and they suggest that the Hh signal is not the whole story and that there is probably a difference in quality between A and P affinities.

Chemotaxis

On the crest of a spiral wave

Drop a pebble into a pool of water and you'll see a series of concentric waves, fanning outwards in ever-increasing circles. Subject slime moulds on a surface to starvation and a similar effect occurs — single-celled *Dictyostelium discoideum* amoebae move chemotactically towards the periodic waves of cyclic AMP that are emitted from the so-called aggregation centre.

Reporting in *Proceedings of the National Academy of Sciences*, Jacques Lauzeral and his colleagues have looked at these cAMP waves and, in particular, at the spiral waves that develop from broken circular waves (*Proc. Natl Acad. Sci. USA* 94, 9153–9158; 1997). By studying the developmental path along which the slime moulds travel, the authors believe that they can now explain how these spiral waves form.

The aggregation centre acts as a 'pacemaker', generating a pulse of cAMP at five- to ten-minute intervals. Under conditions of starvation the amoebae become excitable and, on binding cAMP, an enzyme called adenylate cyclase is activated. This leads to the production of more cAMP, thereby propagating the wave. But, eventually, the amoebae become desensitized to cAMP — they stop binding it, and the wave recedes while the amoebae recover their ability to sense and relay the approaching wave.

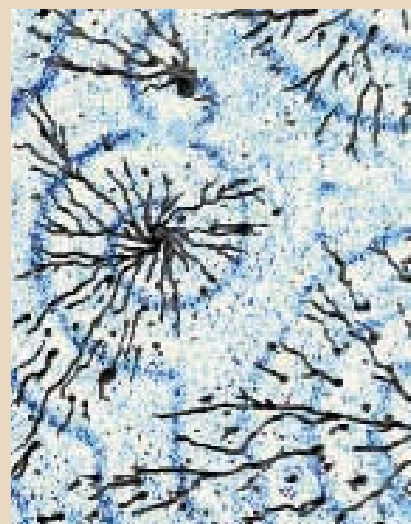
Where do the spiral waves come from? Lauzeral *et al.* claim that the key is desynchronization of cells on the developmental

The interpretation I prefer is therefore that there are at least two contributions to cell affinity — a main one (perhaps graded) related to the positional information that is initiated by and consequent on the Hh signal, and also a difference between A and P that is dependent more directly on Engrailed. □

Peter A. Lawrence is in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

e-mail: pal@mrc-lmb.cam.ac.uk

1. Garcia-Bellido, A. in *Biology of Imaginal Discs* (eds Ursprung, H. & Nothiger, R.) 59–91 (Springer, Heidelberg, 1972).
2. Blair, S. S. & Ralston, A. *Development* **124**, 4053–4063 (1997).
3. Rodriguez, I. & Basler, K. *Nature* **389**, 614–618 (1997).
4. Townes, P. L. & Holtfrete, J. *J. Exp. Zool.* **128**, 53–120 (1955).
5. Morata, G. & Garcia-Bellido, A. *Wilhelm Roux Arch.* **179**, 125–143 (1976).
6. Morata, G. & Lawrence, P. A. *Nature* **255**, 614–617 (1975).
7. Mullor, J. L., Calleja, M., Capdevila, J. & Guerrero, I. *Development* **124**, 1227–1237 (1997).
8. Chen, Y. & Struhl, G. *Cell* **87**, 553–563 (1996).
9. Wright, D. A. & Lawrence, P. A. *Dev. Biol.* **85**, 317–327 (1981).
10. Struhl, G., Barbash, D. & Lawrence, P. A. *Development* **124**, 2155–2165 (1997).



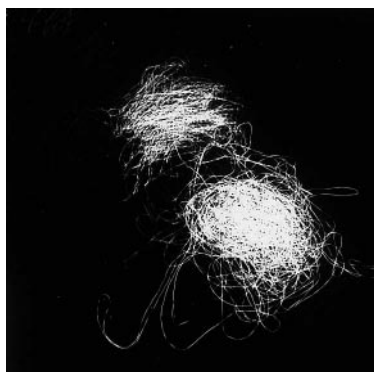
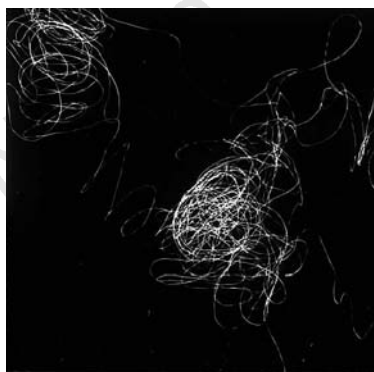
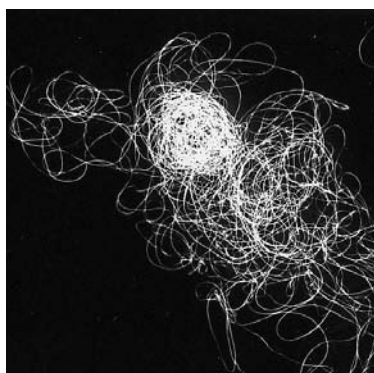
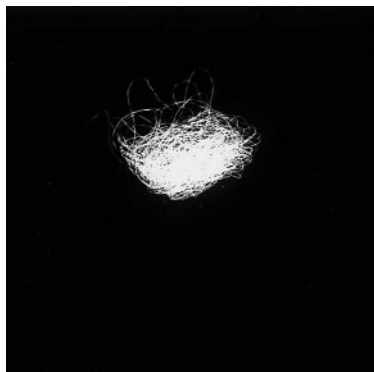
path. Up to 105 cells may aggregate around a single centre, but these cells will not all possess the same amount of adenylate cyclase at any given moment. So, the cells will respond differently to the cAMP pulse — some of them will respond more or less immediately, whereas others may react after a delay. Because the medium is excitable this desynchronized response breaks the concentric circles, leading to the formation of spiral waves (shown in blue in the picture above), the centres of which attract streams of amoebae (black). The upshot is that factors that lead to better synchronization of cells will act as developmental pebbles, generating concentric waves in the cAMP pond.

Alison Mitchell

Avoided object

The negatives of sound

Laquer residue etched from a master recording at Abbey Road studios (mass, easy listening, pop, fugue)



Gamma-ray astronomy

Bursts make new waves

Bohdan Paczyński and Chryssa Kouveliotou

This year there have already been three News and Views articles on γ -ray bursts (GRBs) — violent astronomical events that have been a 30-year puzzle. Why another? The reason is that a meeting on the topic held last month* proved to be historic. The question of 'where' has been resolved in favour of the view that the bursts happen at cosmological rather than galactic distances. Theorists are now concentrating on the 'how' of GRB occurrence. In consequence, we see the beginning of a new era for high-energy astrophysics — GRBs can provide yet another probe of cosmological distances in the early Universe.

The single most important new theoretical concept presented at the meeting was a simple upshot of the notion that GRBs are probably related to the cataclysmic end of massive stars, and therefore that the rate of GRBs is proportional to massive-star formation rate¹. The latter has recently been well established observationally²: it peaked at the redshift $z = 1$ at a rate about ten times higher than it is now.

The consequences are dramatic (D. Hartmann, Clemson Univ.; R. Wijers, Cambridge Univ.). The observed GRB rate is known to follow the so-called Euclidean slope at the bright end: the number of GRBs increases as a -1.5 power of their peak intensity. At the faint end of the distribution, the numbers increase much more slowly, presumably as a result of the cosmological redshift. However, the increase of star formation rate almost exactly compensates for the cosmological dimming, and the predicted counts follow the -1.5 law out to a moderate redshift. This implies that the bursts are considerably farther away, and their intrinsic luminosity is much higher, close to 10^{53} ergs, assuming isotropic emission. A related consequence is that the rate of GRB events per unit volume is reduced by a factor of about 20. So the demand for energy per burst has increased, while at the same time bursts became even rarer than had been commonly thought.

The new era in observations opened in April 1996 with the launch of BeppoSAX, an Italian–Dutch satellite that could detect and locate GRBs fast and accurately, initially using its X-ray Wide Field Cameras and subsequently its X-ray Narrow Field Instruments for a refined position of an X-ray afterglow³. The Wide Field Camera positions were observed from the ground and, for the first time, an optical counterpart was identified in two cases^{4,5}. In the first case (GRB970228), the optical counterpart was embedded in a 'fuzz' reminiscent of a galaxy; in the second

(GRB970508), a lower limit of the distance was spectrally determined⁶, using the powerful Keck telescope, to be $z = 0.835$.

In March and April this year, the Hubble Space Telescope (HST) observed both locations, and confirmed the 'fuzz' of GRB970228 but did not detect a galaxy related to the May event, just a point source^{7,8}. Moreover, using the HST data, a proper motion was claimed⁹ for the February object. The meeting has now cleared up the issue on proper motion: the latest observations, on 4 September with the HST Imaging Spectrograph (A. Fruchter, STScI), confirm that the GRB970228 counterpart is still visible, albeit decaying slowly, but they do not confirm any proper motion — this long temporal baseline makes the null result very strong. The 'fuzz' is again detected at the same level of intensity as in the March/April observations (about 25th magnitude), which adds support to the view that it is a distant dwarf galaxy. This fits with the hypothesis on the host galaxy of GRB970508: a second spectrum¹⁰ of its fading optical counterpart revealed emission lines also corresponding to $z = 0.835$. This result is suggestive of a very compact dwarf galaxy, which is probably related to the GRB and which may become visible as the optical emission from the burst decays further.

Dramatic new results were presented on the radio afterglow following GRB970508 (refs 11, 12; D. Frail, NRAO; G. B. Taylor, NRAO). The afterglow varied rapidly during the first two weeks, but the amplitude of the fluctuations then declined. No proper motion was detected from the source. The simplest explanation of the phenomenon follows the theoretical prediction¹³ that the observed fluctuations are due to scintillation caused by the interstellar medium in our Galaxy. The reduction of the amplitude is explained as a consequence of increase of the angular size of the source, estimated to be initially about 3 micro-arcseconds. Adopting the distance following from the redshift, $z = 0.835$, the expansion rate is moderately superluminal, as predicted by popular models.

But the sceptics at the meeting made the following point. So far, we have followed up, with a compendium of observatories, seven GRBs — only two have been detected in the optical, five have X-ray afterglows associated with them, and only one (GRB970508) has been detected in the radio domain. It is clear that there is no obvious relation between GRB intensity and the presence or intensity of the afterglows in the X-ray, optical or radio domains. The upper limits for any optical emission following GRB970828 are very stringent, and require it to be at least 1,000 times

*Fourth Huntsville Symposium on Gamma-ray Bursts, Huntsville, Alabama, 15–20 September 1997.

weaker than a simple scaling of GRB970228 and GRB970508 suggests (J. van Paradijs, Univs Amsterdam and Alabama, Huntsville).

Fortunately, there may be a simple explanation. The X-ray spectrum obtained with the ASCA satellite (T. Murakami, ISAS) indicated the presence of a large amount of absorbing gas between us and the GRB, much more than expected from our own Galaxy. Assuming a standard gas-to-dust ratio, and assuming that the burst was at a moderate redshift, $z=1$, the implication is that there was so much dust extinction that the optical afterglow should be too dim to be detectable.

The GRB field has now shifted to new wavelengths. The major players today are the X-ray, optical and radio observatories, as well as the γ -ray ones. After a small pause over the summer, BeppoSAX is up and running again. Two other options are now also available, and can provide new targets. First, there is the combination of the Burst And Transient Source Experiment (BATSE), on board the Compton Gamma Ray Observatory, with the Rossi X-ray Timing Explorer (RXTE) and the Ulysses satellite. BATSE detects a burst and defines a relatively small region in the sky where the source should be located. RXTE scans this region for X-ray afterglows; the triangulation of the arrival times of a burst between Ulysses and BATSE provides rings in the sky which further narrow the search area. So far, out of several follow-ups, one X-ray source was detected in this manner. Second, the All Sky Monitor on RXTE recently detected (D. Smith, MIT) two GRBs (GRB970815, GRB970828), providing fast, accurate loca-

tions that have increased our sample of X-ray afterglows.

Finally, the high gas-column density detected in the X-ray spectrum of GRB970828 may be indicative of the association between the GRB and a dense interstellar cloud, possibly a star-forming region. The evidence is inconclusive, but this may be the first hint of a relation between the GRBs and the star-forming regions, as one of us (B.P.) proposed at the Huntsville meeting. However, we may have to wait for the detection of perhaps 100 afterglows before the GRB enigma is finally sorted out (K. Hurley, Univ. California, Berkeley). That should take a lot less than 30 years. □

Bohdan Paczyński is in the Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA.

e-mail: bp@astro.princeton.edu

Chryssa Kouveliotou is with the Universities Space Research Association at NASA/Marshall Space Flight Center, ES-84, Huntsville, Alabama 35812, USA.

e-mail: kouveliotou@msfc.nasa.gov

1. Totani, T. *Astrophys. J.* **486**, L71–L74 (1997).
2. Madau, P. *et al. Mon. Not. R. Astron. Soc.* **283**, 1388–1404 (1996).
3. Costa, E. *et al. Nature* **387**, 783–785 (1997).
4. van Paradijs, J. *et al. Nature* **386**, 686–689 (1997).
5. Bond, H. E. *et al. IAU Circ. No.* 6665 (1997).
6. Metzger, M. R. *et al. Nature* **387**, 878–880 (1997).
7. Sahu, K. *et al. Nature* **387**, 476–478 (1997).
8. Pian, E. *et al. Astrophys. J.* (submitted).
9. Caraveo, P. in *Proc. 4th Compton Symp.*, Williamsburg, VA, April 1997 (in the press).
10. Metzger, M. R. *et al. IAU Circ. No.* 6676 (1997).
11. Frail, D. A., Kulkarni, S. R., Nicastro, L., Feroci, M. & Taylor, G. B. *Nature* **389**, 261–263 (1997).
12. Taylor, G. B., Frail, D. A., Beasley, A. J. & Kulkarni, S. R. *Nature* **389**, 263–265 (1997).
13. Goodman, J. preprint: astro-ph/9706084 (1997).

Signal transduction

Feedback from inhibitory SMADs

Malcolm Whitman

Members of the transforming growth factor- β (TGF- β) superfamily are implicated in biological processes ranging from inhibition of cell proliferation in somatic tissues, to specification of cell fate during embryogenesis. Over the past two years, understanding of the intracellular pathways by which TGF- β signals are mediated has been spurred by studies of the SMAD family of signal transducers¹. Discovered through genetic screens in *Drosophila melanogaster* and *Caenorhabditis elegans*, the biochemistry and cell biology of vertebrate SMADs has implicated them as regulators of the cell-type specific transcriptional responses that are induced by TGF- β ligands. Now, papers in this issue by Imamura *et al.*², Tsuneizumi *et al.*³ and Nakao *et al.*⁴ (pages 622, 627 and 631), along with the work of Hayashi *et al.*⁵ in *Cell*, expose a new aspect of SMADs — some of them inhibit, rather than mediate, TGF- β signalling. These inhibitory SMADs are, themselves, induced by TGF- β

stimulation, suggesting that there is an intracellular negative-feedback loop that regulates TGF- β signals.

Smads 1 to 5 — the canonical signal-transducing SMADs — act downstream of a family of transmembrane serine-threonine kinase receptors for the TGF- β superfamily ligands. Smad1 and, probably, Smad5 mainly act downstream of the bone-morphogenetic-protein (BMP) subset of the TGF- β superfamily. Smad2, and possibly Smad3, act downstream of several other ligands, including TGF- β itself¹ (Fig. 1, overleaf). When their cognate upstream receptors are stimulated by binding of the appropriate ligand, these pathway-specific signal-transducing SMADs are directly phosphorylated at a carboxy-terminal SS(V/M)S consensus site by type I receptors. Phosphorylated (that is, activated), pathway-specific SMADs can each form a stable complex with Smad4, which is found in the signalling pathways of all the different classes of TGF- β superfamily



100 YEARS AGO

Mr. H. Savage Landor, who left England in March last, commissioned by Mr. Harmsworth, the proprietor of the *Daily Mail*, to endeavour to enter the sacred city of Lhasa, in Tibet, has not been successful in his undertaking. News has just been received that a few days after crossing the frontier of Tibet, disguised as a Chinese pilgrim, all except two of Mr. Landor's men abandoned him. In spite of this, Mr. Landor continued on his journey, but eventually he lost all his provisions, and by an act of treachery was made a prisoner by the Tibetans. He was sentenced to be beheaded, but at the last moment the Grand Lama stopped the executioner, and commuted the sentence of decapitation to the torture of the stretching log — a kind of rack upon which Mr. Landor was chained for eight days — after which he was released. Mr. Landor has now returned to India, suffering from the effects of the torture to which he was subjected, and which he half anticipated before he set out upon his hazardous journey. From *Nature* 7 October 1897.

50 YEARS AGO

From the obituary of Hans Fischer, "to whom we owe most of our knowledge of the chemistry of hæmin, chlorophyll and the bile-pigments". Fischer, who won a Nobel prize in 1930, died on 31 March 1945, but *Nature* only became aware of his death several months later.

There was an atmosphere of unusual jollity in [his] laboratory, directed in its legitimate expression by Herr Paulus, the store-keeper, and controlled, in its more outrageous excesses, by Fischer's tact and humour. A fledgling 'doctor', returning from his oral examination — a formality — found his bench littered with the starting materials for a celebration and a large blackboard, on which a cartoon and some lines of doggerel reminded him that he was mortal. More rarely, there were occasions in the cellars, and at Christmas the supply of 5-litre flasks was exhausted as all undertook the preparation of 'christmas-pyrrole' by a process supposed to render denatured alcohol potable.

● We regret also to announce the death of Prof. Max Planck, For.Mem.R.S., on October 4, aged eighty-nine. From *Nature* 11 October 1947.

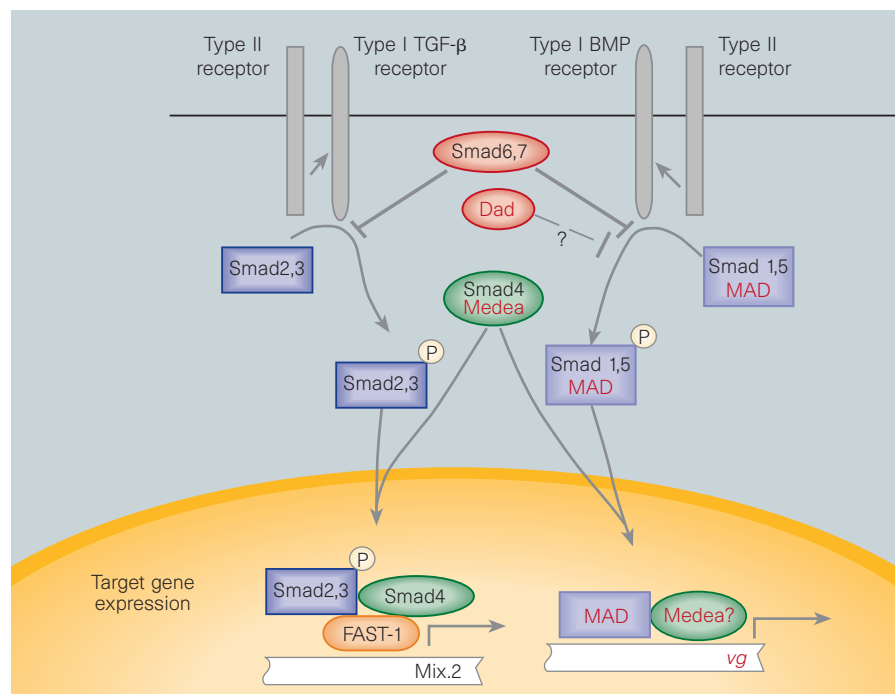


Figure 1 SMADs involved in the transmission and inhibition of transforming growth factor- β (TGF- β) superfamily signals. Ligand-binding type II receptors associate with, phosphorylate and activate type I receptors (for TGF- β or bone morphogenetic protein; BMP). The type I receptors phosphorylate pathway-specific SMADs (Smads 1, 2, 3 and 5). These activated SMADs then associate with Smad4, and translocate to the nucleus where they may regulate transcription either by associating with nuclear transcription factors, as in the case of the *Mix.2* promoter⁹, or by binding directly to DNA, as in the case of the *vestigial* (*vg*) promoter¹⁰. The inhibitory SMADs described by Imamura *et al.*², Tsuneizumi *et al.*³, Nakao *et al.*⁴ and Hayashi *et al.*⁵ target the first step in the intracellular transduction pathway — type I receptor phosphorylation of pathway-specific SMADs. For simplicity, additional homomeric interactions among receptors and SMADs are not shown. *Drosophila* proteins are shown in red. Medea is a *Drosophila* homologue of Smad4 (ref. 11), and the association of Medea with the *vestigial* promoter is hypothetical.

ligands. This complex then translocates to the nucleus, where it regulates transcriptional responses to TGF- β s¹. The pathway-specific SMADs (Smads 1, 2, 3 and 5) all associate with type I TGF- β -superfamily receptors and are phosphorylated at the conserved carboxy-terminal site; Smad4 functions through association with these pathway-specific SMADs.

Now, Imamura *et al.*², Nakao *et al.*⁴ and Hayashi *et al.*⁵ report that Smad6 and Smad7 inhibit the vertebrate SMAD-based signalling pathway described above. Moreover, Tsuneizumi *et al.*³ show that, in a homologous signalling pathway in *Drosophila*, the product of the *Daughters against decapentaplegic* (*Dad*) gene is an inhibitory SMAD. Overexpression of these newly discovered SMADs inhibits intracellular signalling by members of the TGF- β superfamily. Smad6 and Smad7 can inhibit both TGF- β and BMP signalling in cultured cells or in frog embryos^{2,4,5}. The *Dad* gene inhibits patterning by *decapentaplegic*, a *Drosophila* BMP homologue, in the *Drosophila* imaginal wing disk³. Like Smad4, these inhibitory SMADs lack the carboxy-terminal phosphorylation site that is found in each of the pathway-specific SMADs. But unlike Smad4, Smads 6

and 7 do interact with type I receptors — and they probably interact more stably than do the pathway-specific SMADs. Presumably, this is because Smad6 and Smad7 lack the carboxy-terminal phosphorylation site, so they are not phosphorylated and released from activated receptor kinase.

So, competitive inhibition of the SMAD binding site on the type I receptor kinase may explain how Smad6 and Smad7 inhibit signalling by members of the TGF- β superfamily. Overexpression of Smad6 or Smad7 inhibits ligand-stimulated phosphorylation of pathway-specific SMADs^{2,4,5}, indicating that these inhibitory SMADs can block signalling at this initial step in the intracellular transduction pathway. In theory, an inhibitory SMAD might work by sequestering signalling SMADs or downstream effectors, as well as by tying up receptors. Intriguingly, Topper *et al.*⁶ have identified a truncated form of Smad6 that is co-induced with Smad7 by shear stress in vascular endothelial cells. The truncated Smad6 stably associates with Smad1, Smad2 and Smad4, as well as with Smad7. So, some inhibitory SMADs might also act by sequestration of signalling SMADs.

Defining the specificities of inhibitory SMADs with respect to upstream ligands

and receptors, and downstream signalling SMADs, will be critical for understanding their biological functions. Overexpression of Smad7 inhibits phosphorylation of Smad2 and Smad3 by activated type I TGF- β receptor, and prevents phosphorylation of Smad1 by activated BMP type I receptors. Smad7 could, therefore, be a generalized inhibitor of signalling by the TGF- β superfamily⁴. Smad6 is more complicated — it inhibits phosphorylation of Smad2 but not Smad3. Moreover, it inhibits phosphorylation of Smad1 by the BMP type IB receptor, but not by the BMP type IA receptor. Analysing the specificity of inhibitors by overexpression is a tricky business, because inhibition *in vivo* would be expected to be a function of the endogenous concentrations of type I receptors, signalling SMADs and inhibitory SMADs. But inhibitory SMADs that distinguish between different ligands or receptor subtypes of the TGF- β superfamily ligands would provide considerable flexibility to a cell's repertoire of responses to TGF- β stimuli.

The question surrounding specificity of the inhibitory SMADs is closely tied to understanding their role in feedback-inhibition of TGF- β signals. If each inhibitory SMAD targets the full range of type I receptors in the TGF- β superfamily, feedback-inhibition from one ligand will reduce responsiveness to the larger superfamily of ligands. Alternatively, if inhibitory SMADs are specific for distinct receptors or signalling SMADs, they may provide a mechanism by which cells can independently regulate either sensitivity to distinct ligands or specific downstream responses. The timing for induction of inhibitory SMADs may also be an important aspect of feedback control — Smad7 is induced much more rapidly than the product of the TGF- β -responsive *PAI-1* gene. So we need to know how inhibitory SMADs modulate not only the magnitude but also the duration of a TGF- β signal.

Although SMADs have become a focus of research, other signalling pathways have been implicated in TGF- β -mediated signal transduction⁷. If inhibitory SMADs specifically target the SMAD component, it is possible that feedback-regulation by inhibitory SMADs does not downregulate TGF- β signalling entirely. The effects of inhibitory SMADs on a variety of transcriptional and phenotypic responses to TGF- β s are reported in the new papers, but a more exhaustive examination of biological responses to TGF- β will be necessary to determine whether inhibitory SMADs block some or all components of TGF- β signalling.

Ligands of the TGF- β superfamily are morphogens in both vertebrate and invertebrate⁸ embryos. This suggests that the spatially and temporally complex regulation of responses to TGF- β s across developing tissues is fundamental to embryonic patterning. Now, the discovery that inhibitory

SMADs are feedback regulators of TGF- β signalling opens up the question of how the expression and regulation of these molecules participates in the patterning of TGF- β responses in developing tissues. □

Malcolm Whitman is in the Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA.

e-mail: mwhitman@warren.med.harvard.edu

1. Massagué, J., Hata, A. & Liu, F. *Trends Cell Biol.* 7, 187–192 (1997).

2. Imamura, T. *et al. Nature* **389**, 622–626 (1997).
3. Tsuneizumi, K. *et al. Nature* **389**, 627–631 (1997).
4. Nakao, A. *et al. Nature* **389**, 631–635 (1997).
5. Hayashi, H. *et al. Cell* **89**, 1165–1174 (1997).
6. Topper, J. N. *et al. Proc. Natl Acad. Sci. USA* **94**, 9314–9319 (1997).
7. Alevizopoulos, A. & Mermod, N. *Bioessays* **19**, 581–591 (1997).
8. Neumann, C. & Cohen, S. *Bioessays* **19**, 721–729 (1997).
9. Chen, X. *et al. Nature* **389**, 85–89 (1997).
10. Kim, J., Johnson, K., Chen, H. J., Carroll, S. & Laughon, A. *Nature* **388**, 304–308 (1997).
11. Raftery, L. A., Twombly, V., Wharton, K. & Gelbart, W. *Genetics* **139**, 241–254 (1995).

Telomerase

Cancer and the knockout mouse

David Wynford-Thomas and David Kipling

Telomeres are specialized structures at the ends of linear chromosomes. They usually consist of tandem arrays of a short DNA sequence (TTAGGG in vertebrates) and associated proteins, and they are thought to have at least two essential functions¹. First, they stop natural chromosome ends behaving as random breaks, which might otherwise generate inter-chromosomal fusions and activate DNA-damage-induced cell-cycle arrest. Second, they provide the structural basis for solving the end-replication problem² — the inability of DNA polymerases to completely replicate the end of a DNA duplex (see box).

In the germ line, telomeres are maintained by the compensatory addition of TTAGGG repeats to chromosome ends. These repeats are synthesized by an enzyme called telomerase (Fig. 1, overleaf), and a lack of telomerase activity — such as occurs physiologically in most adult human somatic cells — leads to progressive telomere shortening with every cell cycle³. If such cells are forced to grow for long enough, they eventually lose telomere function leading to end-to-end chromosome fusions and cell death. What, then, might be the effect of a constitutional (germline) deficiency of telomerase? An intriguing (although still partial) answer is provided in the latest issue of *Cell*, where Blasco *et al.*⁴ report their studies of a telomerase-deficient mouse.

Telomerase contains an essential RNA component that provides the template for the specific synthesis of TTAGGG repeats. Blasco *et al.* used transgenic technology to create a germline deletion of the *mTR* gene, which encodes the RNA component of mouse telomerase. The authors bred homozygous, null *mTR*^{-/-} mice which turned out to be both viable and fertile, despite having no detectable telomerase activity. Moreover, the mice have now been maintained for six generations (G1–G6). The absence of any initial phenotype is striking. So is the fact that the next few generations have also shown no symptoms, suggesting that the telomerase inhibitors that have been envisaged for cancer therapy will not have

any acute toxicity. But telomere shortening is occurring: by the latest generation studied (G6), there is clear evidence of telomere erosion, with around five per cent of chromosomes in embryonic fibroblasts lacking detectable TTAGGG, accompanied by increasingly frequent chromosome fusions.

This 'delay' in manifestation of the phenotype almost certainly reflects the unusual long telomere-repeat arrays in the germ line of this mouse species. These repeats are not exhausted until many generations have elapsed — a human 'knockout' would be predicted to show a much more rapid onset. Yet, despite reaching what seems to be a critical state of erosion, cells from generation G6 do not show an obvious reduction in growth

capacity in culture, or any reduction in the ability to generate tumours when oncogene-transformed cells are injected into nude (immunodeficient) mice.

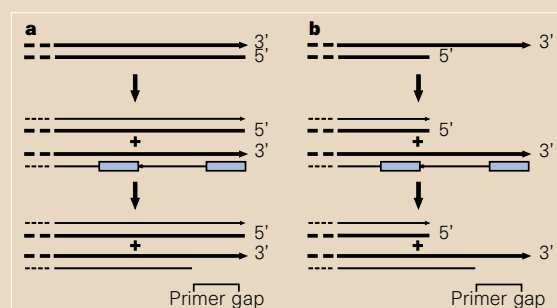
It has been suggested that progressive erosion of telomeres in somatic cells (which have physiologically repressed telomerase) represents a barrier to indefinite cell proliferation. Tumour cells are thought to overcome this block by reactivating telomerase. But the apparently undiminished tumorigenic ability of oncogene-transformed *mTR*^{-/-} cells suggests that telomerase may not be necessary for tumorigenesis. Data showing that, in mice, telomerase can be up-regulated at an early stage in tumour development⁵ — before any significant telomere erosion has occurred — leads the authors to hint that the current dogma may be wrong. In other words, telomerase may be nothing more than a passive bystander, rather than facilitating tumour growth⁴. If so, this would have considerable implications for telomerase as a target for cancer therapy (although it does not necessarily detract from its value as a diagnostic marker).

There are, however, good reasons for exercising caution when extrapolating from the mouse model. The main limitation is that the phenotype is probably not yet completely developed. Although telomere failure is occurring at G6, most of the chromosomes retain detectable TTAGGG. The very fact that viable G6 animals exist suggests

The end-replication problem

For lagging-strand DNA replication, short RNA primers (blue) are made by RNA primase. These are then extended by DNA polymerase to form Okazaki fragments.

When these RNA primers are removed, there is no way to synthesize lagging-strand sequence that is complementary to the small region at the end of the chromosome (which is at least as large as an RNA primer). So, with continuing cell division, sequence is lost from the ends of linear chromosomes. Some blunt-ended daughter molecules are produced by this scheme, irrespective of whether the starting terminus is blunt-ended (a) or has a



3' extension (b).

Why are such natural blunt ends not recognized as DNA damage? One possibility is that this is because unnatural ends have a slightly different chemical structure; for example, radiation-induced blunt ends can have terminal 3' phosphoglycolate residues whereas physiological ends do not. Human cells may have an additional

mechanism to ensure that natural chromosome ends are even more different. This involves a degradative pathway which, even in telomerase-deficient human cells, results in 3' extensions for most chromosome ends⁹. This may be analogous to a degradative pathway that has been described in budding yeast, involving Cdc13p and other proteins¹⁰. **D.W.-T. & D.K.**

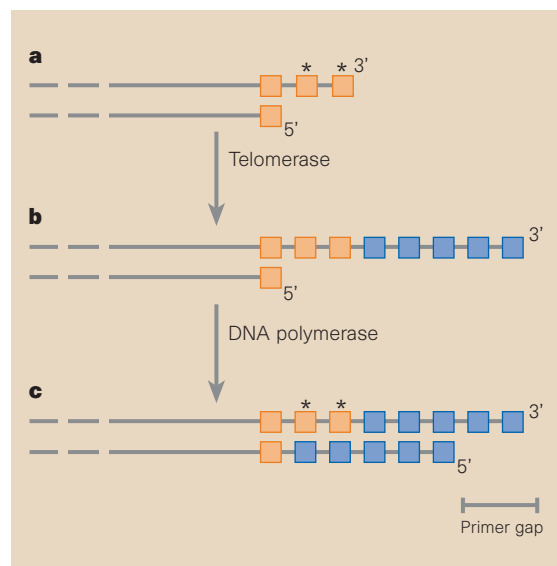


Figure 1 Synthesis of telomeric DNA by telomerase. **a**, A short region of the upper strand of the telomere (composed of TTAGGG repeats) cannot be copied by normal DNA replication (see box) and is indicated by asterisks. **b**, Telomerase circumvents this by synthesizing new TTAGGG repeats (blue), allowing the ends to be copied by conventional DNA replication (**c**). Mouse telomerase carries an RNA molecule that acts as a template for synthesis of the TTAGGG sequence. Blasco *et al.*⁴ have deleted the *mTR* gene (which encodes this RNA molecule) in the mouse germ line. The resultant knockout mice have no telomerase activity, and their telomeres become shorter with continuing cell division.

that interference with chromosome function may need to be more widespread to compromise cell proliferation severely. It seems improbable that the observed telomere erosion could continue unabated without eventually stopping growth. For example, viral-oncogene-transformed human cells eventually undergo 'crisis' (cell death) during continuous culture, unless telomere erosion is halted by activation of either telomerase³ or of an alternative pathway⁶. A remarkable counterpart *in vivo* may be the spontaneous tumour regression that occurs in a subset of childhood neuroblastomas (so-called stage IV-S), which apparently fail to up-regulate telomerase⁷. So, if telomere erosion continues, oncogene-transformed *mTR*^{-/-} cells might be expected, eventually, to lose their ability to generate tumours, once they have undergone enough cell divisions.

It is also risky to extrapolate from the mouse with regard to the timing of telomerase up-regulation in tumour development — not least because telomerase seems to be more tightly regulated in man. Many (but not all) human cancers possess very short telomeres, consistent with the classic model of telomere erosion followed by selection for telomerase reactivation during an *in vivo* equivalent of crisis. Of course, some tumours may acquire telomerase ahead of critical telomere erosion; for example, if telomerase is reactivated as part of a 'package' of changes in gene expression that occurs after some other genetic event (the 'co-selection' hypothesis^{4,8}). But such anticipatory up-regulation does not mean that telomerase is not required at a later stage, so this should not dampen enthusiasm for anti-telomerase therapies.

We may also need to revise our ideas about how cells distinguish between unnatural genomic breaks such as those caused by ionizing radiation (which usually lead to immediate cell-cycle arrest or programmed

cell death), and natural chromosome ends (which do not). It has been assumed that in mammalian cells, specific terminal DNA sequences and associated proteins somehow 'shield' the natural end of the chromosome and prevent it from activating genome-damage-monitoring systems. The later-generation *mTR*^{-/-} cells show an apparent loss of telomere-protective function (as judged by end-to-end fusions in metaphase chromosome preparations). But the fact that such cells have reached metaphase suggests that they have passed through at least one damage-sensing checkpoint (at G2-M), despite having critically eroded telomeres.

If the absence of telomere-specific DNA sequences and proteins does not activate cell-cycle arrest, then what does? It is possible that the physical structure of the DNA-end itself is differentially recognized — a break caused by ionizing radiation is not only physiologically abnormal, but it is chemically very different from a natural chromosome terminus formed by DNA replication.

We now await with interest the phenotype of later generations of these mice (if they remain fertile), and later passages of cell cultures derived from them. Whatever the outcome, it is certain that interest in telomeres has not come to an end. □

David Wynford-Thomas and David Kipling are in the Department of Pathology, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN, UK.

- Kipling, D. *The Telomere* (Oxford Univ. Press, 1995).
- Olovnikov, A. M. *Dokl. Akad. Nauk* **201**, 1496–1499 (1971).
- Counter, C. M. *et al.* *EMBO J.* **11**, 1921–1929 (1992).
- Blasco, M. A. *et al.* *Cell* **91**, 25–34 (1997).
- Broccoli, D., Godley, L. A., Donehower, L. A., Varmus, H. E. & de Lange, T. *Mol. Cell. Biol.* **16**, 3765–3772 (1996).
- Bryan, T. M. & Reddel, R. R. *Eur. J. Cancer* **33**, 767–773 (1997).
- Hiyama, E. *et al.* *Nature Med.* **1**, 249–255 (1995).
- Kipling, D. *Eur. J. Cancer* **33**, 792–800 (1997).
- Makarov, V. L., Hirose, Y. & Langmore, J. P. *Cell* **88**, 657–666 (1997).
- Lin, J.-J. & Zakian, V. A. *Proc. Natl Acad. Sci. USA* **93**, 13760–13765 (1996).

Daedalus

Smooth ascent

A solid-fuel rocket is simply an enormous firework. Once lit, it cannot be stopped or controlled. Furthermore, it burns in massively turbulent flow, generating extreme vibration. Everything that has to fly in a rocket must pass a searching vibration test. Daedalus is now devising a less vibrant way to climb into space.

A typical solid fuel is a mixture of aluminium powder and ammonium perchlorate, formed into a resilient solid by a polymeric binder. DREADCO chemists are replacing the aluminium powder by fine aluminium wire, and the binder by a water-based gel. The product is just damp enough so that ignition cannot spread through it spontaneously; but each element can be ignited by passing a current through the local aluminium wire. The charge will be fired controllably, element by element, by passing a programmed sequence of current pulses through the wires.

In this simple form, the idea is clearly impractical. A big rocket motor could have millions of ignitable elements, far too many for each to have its own electrical leads. But Daedalus recalls the old flashbar units for small cameras, which carried ten separate flash-bulbs. As each bulb fired, it carbonized and caused a lead to the next bulb to become conducting. Each time the shutter was depressed, the trigger current set off the next live bulb. This trick, suitably developed, would allow thousands of elements in the motor to be set off controllably by a single set of leads. A few thousand sets of such leads could command the complete motor.

The resulting rocket will be a masterpiece of smooth power. Its thrust will be controlled from instant to instant by the rate at which its elements are fired. Its vibration will be damped by fast-acting acceleration sensors feeding its control computer, which will stagger the firing sequence to the next elements so as to cancel or minimize the instantaneous fluctuations of thrust. If required, the whole motor can be shut off instantly, and started again later; once in space, it can be operated on command at very low thrust for orientation or course correction.

The technology could even be adapted for other purposes, such as blasting, demolition and explosive forming. Instead of many explosives for different jobs, one all-purpose DREADCO product will do. It can be programmed to act as a propellant, a low, medium or high explosive, or even a subtle combination of each.

David Jones

Systemic signalling in gene silencing

Gene silencing in plants is a genetic control mechanism implicated in virus resistance^{1,2}, genome maintenance³ and developmental control⁴. We describe here our recent discovery that there is a systemic signal that can mediate gene silencing. From the gene-specificity of the systemic silencing, we infer that the signal molecule is likely to be a nucleic acid.

We studied *Nicotiana benthamiana* plants carrying a jellyfish green fluorescent protein (GFP) transgene⁵. We infiltrated leaves with strains of *Agrobacterium tumefaciens* carrying a GFP reporter gene. The bacterium contains a tumour-inducing (Ti) plasmid which can be modified in the transferred DNA (T-DNA) region to carry foreign genes into plant nuclear DNA.

Two days after infiltration, there was strong green fluorescence in the infiltrated region, indicating transfer of the T-DNA from the bacterium into the plant cell and expression of the GFP reporter gene (Fig. 1a). This strong fluorescence was superimposed over a weaker background fluorescence from the GFP found naturally in the transgenic plants. However, at the edge of the infiltrated zone there was a thin line of tissue in which the GFP fluorescence was absent, indicating that expression of the GFP transgene had been silenced (Fig. 1a).

Although the zone of GFP silencing did

not spread further within the infiltrated leaf, by 18 days after infiltration there was silencing of the GFP transgene in the upper leaves of the infiltrated plant (Fig. 1b–d). In leaves of the axillary shoots (Fig. 1b) and in some of the uppermost leaves (Fig. 1d) there was complete silencing of the GFP transgene. The GFP transgene was not silenced if the Ti plasmid did not have a GFP coding sequence; if the cultures were prepared in the absence of acetosyringone as an inducer of DNA transfer into plant cells⁶; or if the leaves were infiltrated with water (Fig. 1d). We conclude that the systemic silencing of the GFP transgene is a sequence-specific effect, following transient expression of the GFP reporter gene from the bacterium in the transgenic plants.

In subsequent tests, we found that systemic silencing of the GFP transgene was associated with reduced levels of GFP messenger RNA. We also found that an RNA virus vector, potato virus X carrying the GFP gene (PVX-GFP)⁷, failed to accumulate when inoculated into the GFP-silenced leaves. This resistance to PVX-GFP was sequence-specific because PVX with an insert of the GUS reporter gene accumulated to a high level. It is likely that this sequence-specific virus resistance and the systemic silencing are the result of the same mechanism targeted against the GFP

coding sequence of the viral and transgene RNAs.

Using several approaches, we failed to detect either *A. tumefaciens* cells or T-DNA in the systemic tissue, indicating that the silencing was caused by a signal molecule that moves out of the infiltrated leaf. The signal was produced when the GFP reporter gene was transferred to the cells of the GFP transgenic plant and is targeted against GFP RNA, indicating that the GFP, or the corresponding DNA or RNA, might be a component of the signal.

Of these, GFP is the least plausible candidate, because there is no mechanism known to us that explains how it could move systemically and specifically target the GFP RNAs. However, a nucleic acid-based signal could mediate sequence-specific gene silencing by forming a base-paired or triple helical structure with the target RNA. Moreover, a nucleic acid could move in the plant using the channels involved in virus or viroid movement.

Previous analyses have described non-clonal mosaics of gene silencing^{8,9} and inferred that an extracellular signal is involved, but these studies did not suggest that the signal would have the potential to move systemically. The systemic signal may be related to recent findings that gene silencing is associated with induced, natural defence against viruses^{1,2}. The signal could move in the plant ahead of the inducing virus, so that antiviral gene silencing could delay spread of the infection front.

We are currently attempting to characterize the systemic signal of GFP gene silencing. Meanwhile, the system described here may provide an experimental system for testing the function of cloned plant genes and for dissecting the mechanism of gene silencing.

Olivier Voinnet

David C. Baulcombe

The Sainsbury Laboratory,
John Innes Centre, Colney,
Norwich NR4 7UH, UK
e-mail: baulcombe@bbsrc.ac.uk

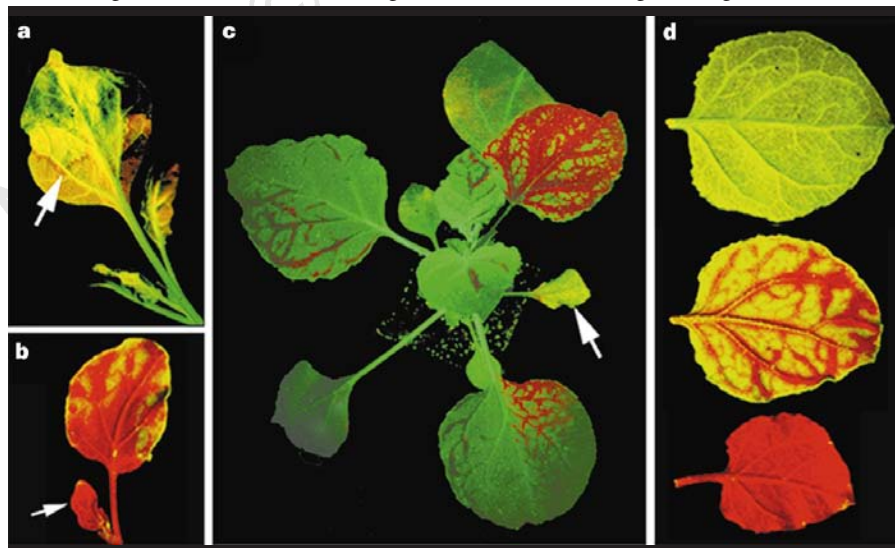


Figure 1 Expression of GFP in transgenic *N. benthamiana*. Under ultraviolet illumination, GFP appears green or yellow. In the absence of GFP, the tissue appears red owing to the fluorescence of chlorophyll. **a**, The leaf of GFP transgenic *N. benthamiana* infiltrated 2 days previously with *A. tumefaciens* carrying a GFP reporter gene. Arrow indicates the GFP-silenced boundary between the distal infiltrated zone (strong green fluorescence) and the proximal zone (weak green fluorescence due to the GFP transgene). **b**, Leaf and axillary shoot (arrow) of the upper part of a plant infiltrated 18 days previously, showing residual GFP in diffuse regions of the leaf and the extreme distal region of the axillary shoot. **c**, Intact GFP transgenic plant infiltrated 18 days previously in a lower leaf (arrow) showing the progression of GFP-silencing. **d**, Upper leaves emerging from the main stem of GFP plants infiltrated 18 days previously with water (top), or with the GFP strain of *A. tumefaciens* (middle and bottom).

1. Covey, S. N., Al-Kaff, N. S., Langara, A. & Turner, D. S. *Nature* **385**, 781–782 (1997).
2. Ratcliff, F., Harrison, B. D. & Baulcombe, D. C. *Science* **276**, 1558–1560 (1997).
3. Assaad, F. F., Tucker, K. L. & Signer, E. R. *Plant Mol. Biol.* **22**, 1067–1085 (1993).
4. Boerjan, W., Bauw, G., van Montagu, M. & Inzé, D. *Plant Cell* **6**, 1401–1414 (1994).
5. Haseloff, J., Siemering, K. R., Prasher, D. C. & Hodge, S. *Proc. Natl Acad. Sci. USA* **94**, 2122–2127 (1997).
6. Ream, W. *Annu. Rev. Phytopathol.* **27**, 583–618 (1989).
7. Baulcombe, D. C., Chapman, S. N. & Santa Cruz, S. *Plant J.* **7**, 1045–1053 (1995).
8. Jorgensen, R. A. *Science* **268**, 686–691 (1995).
9. Kunz, C., Hanspeter, S., Stam, M., Kooter, J. M. & Meins, F. J. *Plant J.* **10**, 437–450 (1996).

A pandemic warning?

Introduction of new influenza type-A viruses, carrying different combinations of the viral envelope glycoproteins haemagglutinin (H) and neuraminidase (N), have led to three major pandemics of influenza in humans this century. Phylogenetic evidence suggests that these viruses have originated from avian influenza A viruses, either unchanged or after reassortment with human influenza A viruses. In aquatic birds, all of the known H and N antigenic varieties (15 varieties carry H, nine carry N envelope glycoproteins) apparently circulate in a genetically conserved fashion. Viruses carrying the H1N1, H2N2 and H3N2 combinations were responsible for the Spanish flu of 1918, the Asian flu in 1957 and Hong Kong flu in 1968, respectively¹. An influenza A virus of the H5N1 subtype has now been identified in a human patient, raising discussions about its potential to spark a new human influenza pandemic.

On 21 May 1997, the fifth day of his hospitalization, a three-year-old boy from Hong Kong died in an intensive-care unit in Hong Kong, with a final diagnosis of Reye syndrome, acute influenza pneumonia and respiratory distress syndrome (ARDS). No

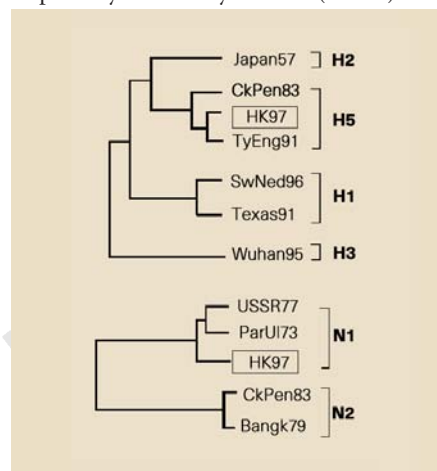


Figure 1 Phylogenetic relationship of the H and N genes from the human virus isolate A/Hongkong/156/97 (HK97) and the corresponding genes from selected influenza A viruses. Japan57, A/Japan/305/57 (H2N2); CkPen83, A/chick/Pennsylvania (H5N2); TyEng91, A/turkey/England/5092/91 (H5N1); SwNed96, A/swine/Netherlands/609/96 (H1N1); Texas91, A/Texas/36/91 (H1N1); Wuhan95, A/Wuhan/359/95 (H3N2); USSR77, A/USSR/90/77 (H1N1); ParUI73, A/parrot/Ulster/73 (H7N1); Bangk79, A/Bangkok/1/79 (H3N2). Sequences were from GenBank or nucleotide sequence analysis using the Dye Deoxy Terminator kit (Applied Biosystems, Foster City, CA). Alignment was done with 827 nucleotides (38–865) of the H genes and 200 nucleotides (1–200) of the N genes by PILEUP (GCG, Madison, WI). Phylogenetic trees were constructed in Phylip (Felsenstein, Seattle, WA) using DNAPARS and DRAWGRAM.

indications of other underlying disease, including immunodeficiency or cardio-pulmonary disease, were observed. From a tracheal aspirate, we isolated an influenza virus in MDCK and LLC cells. We were unable to grow any pathogenic bacteria from respiratory specimens. In haemagglutination inhibition assays, the virus did not react with ferret antisera to recent isolates of human and swine subtypes. By complement fixation tests we found it to be an influenza A virus, confirmed by an indirect immunofluorescence assay on cells from the original tracheal aspirate.

Haemagglutination inhibition assays using antisera to 14 H subtypes showed that the isolate was an H5 influenza A virus. Biochemical neuraminidase inhibition tests, using antisera to nine N subtypes, indicated that the neuraminidase was of the N1 subtype. Nucleotide sequence analyses of parts of the H and N genes of the virus allowed a phylogenetic comparison with other influenza viruses. Our analyses clearly confirmed that the virus is of the H5N1 subtype (Fig. 1). The H5N1 virus was indeed isolated from the sample of the child, and could not be attributed to a laboratory cross-contamination, as there was no H5N1 virus present in the human diagnostic laboratory. Furthermore, we reisolated the same virus from the original tracheal aspirate specimen.

The contribution of the influenza A H5N1 virus infection to the child's disease, eventually leading to his death, is not yet clear. But the virus identification is important as it is the first documented isolation of an influenza A virus of this subtype from humans. Subtype H5 influenza A viruses can cause lethal avian influenza (fowl plague), a disease that may decimate flocks of domestic poultry. These viruses have so far not been identified in pigs. We feel that the identification of the H5N1 influenza A virus and its presently unknown pandemic potential, should be the basis for an intensive monitoring of the epidemiology and the clinical manifestation of infection with this virus by the international World Health Organization influenza surveillance network.

J. C. de Jong, E. C. J. Claas

A. D. M. E. Osterhaus

National Influenza Centre of The Netherlands, Erasmus University Rotterdam and National Institute of Public Health and the Environment, Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands

R. G. Webster

Department of Virology and Molecular Biology, Saint Jude Children's Research Hospital, 332 North Lauderdale Street, Memphis, Tennessee 38105-2794, USA

W. L. Lim

Government Virus Unit, Queen Mary Hospital, Department of Health, The Hong Kong Special Administrative Region

1. Webster, R. G. *et al.* *Microb. Rev.* **56**, 152–179 (1992).

Conical beams from open nanotubes

Electron guns are indispensable devices that are widely used in household and industrial appliances. Field electron-emitting sources (which emit electrons by tunnelling effects in electric fields), with their small size, small energy spread, high current density and no requirement for heat, have distinct advantages over thermionic emitters. We have made a field electron emitter from hollow, open-ended carbon nanotubes.

The most commonly used material for field emitters is tungsten, which operates only under ultra-high vacuum conditions. It is for this reason that field electron emitters have not been widely adopted commercially. Carbon nanotubes have intrinsically suitable properties for acting as field emitters: sharp tips with a nanometre-scale radius of curvature, high mechanical stiffness, chemical inertness and electrical conductivity. Indeed, field emission from an individual multi-walled nanotube (MWNT)¹ as well as from assemblies of MWNTs^{2,3} has been shown under conventional high-vacuum conditions. But there have so far been no microscopic studies of the geometrical or atomic structure of the emitting regions of nanotube tips.

Using field-emission microscopy (FEM)⁴, we have now observed emission patterns from two kinds of MWNT — ordinary MWNTs with closed caps and open-ended MWNTs (Fig. 1). We retrieved the first type from the core of a cathode deposit produced by a carbon arc. We produced the open-ended type from the first using a purification process, by which graphite grains and nanoparticles were removed by oxidation with the aid of CuCl₂ intercalation⁵. We

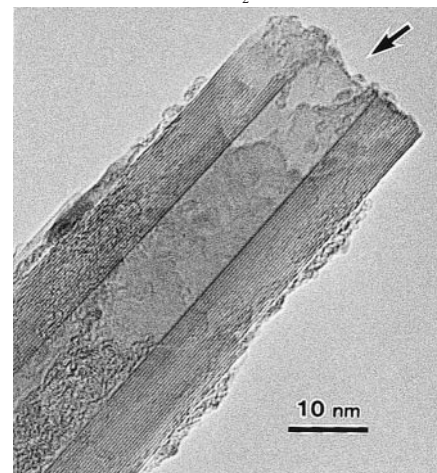


Figure 1 Transmission electron microscope image of the open end of a multi-walled carbon nanotube. The exposed cavity is indicated by an arrow. The cap of the nanotube was removed using a purification process that included oxidation.

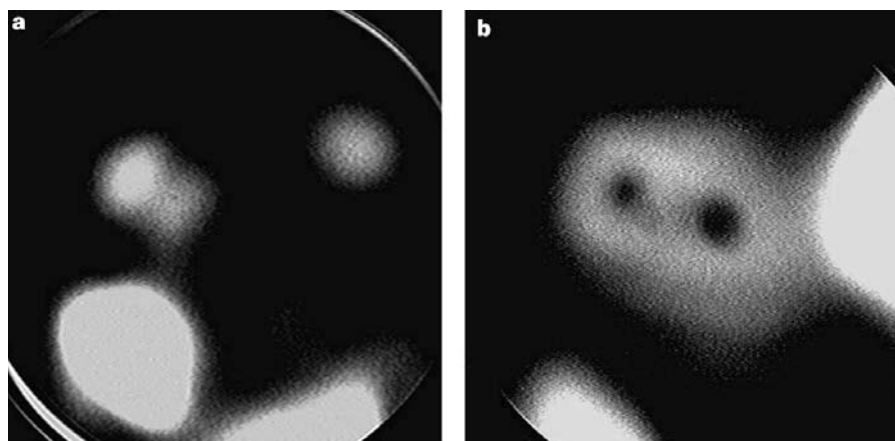


Figure 2 Field-emission patterns from multi-walled carbon nanotubes. **a**, Field-emission pattern from closed-capped MWNTs. **b**, Field-emission pattern from open-ended MWNTs. The distance between the nanotube emitter (cathode) and the microchannel plate (anode) was 60 mm. The electric potential of the emitter against the grounded anode was -500 V for the closed tubes and -320 V for the open tubes.

purified MWNTs in the form of a black, thin 'mat' (a flake with a thickness of a few hundredths of a millimetre).

To make an electron emitter from capped MWNTs, we picked up a needle-like fragment (~ 0.1 mm diameter, 1–2 mm long) from the cathode deposit, and fixed it to a hairpin-shaped tungsten or nichrome wire (diameter 0.3 mm) using carbon paste. For the purified MWNTs, we cut out a thin thread (less than 0.1 mm wide, 1–2 mm long) with a sharp tip from the mat of MWNTs, using a razor, and attached a wire as before. We positioned the MWNT 60 mm in front of a microchannel plate, which had an effective diameter of 42 mm, and placed a fluorescent screen just behind the plate. The working pressure of the vacuum chamber for FEM was typically 2×10^{-9} torr, and the emitter tip was cooled to near liquid-nitrogen temperature. Emission patterns were recorded with a charge-coupled device (CCD) camera (Hamamatsu C5985).

We saw no inner structure in the capped MWNTs (Fig. 2a), but the open tubes showed a peculiar pattern of circular bright rings (Fig. 2b). In Fig. 2b there are two rings in contact with each other, corresponding to adjacent open tubes. A black spot in the central region — the absence of electrons in the core of a beam — corresponds to an exposed cavity. Diameters of inner black spots measured on the screen (2–3.5 mm) and the inner diameters of MWNTs (typically 5–10 nm) indicate an approximate magnification of one million for our FEM. As the electrons emitted from the tip are radially accelerated toward the screen, the beam from an open tube forms a hollow cone. The apex angles of the outer and inner cones of the hollow beam are about 0.2 and 0.05 radians, respectively. Using open-ended nanotubes we can produce a unique tubular configuration of the electron beam.

Rinzler *et al.*¹ claimed to have observed field emission from a linear carbon chain

pulled out from the open edge of a nanotube. Contrary to their report, we saw no sharp contrast corresponding to the atomic chain in our emission patterns. Electron emission seems to occur from the circular edges of the graphite layers of a nanotube.

Yahachi Saito, Koji Hamaguchi, Koichi Hata

*Department of Electrical and Electronic Engineering,
Mie University, Tsu 514, Japan
e-mail: saito@is.elec.mie-u.ac.jp*

Kunio Uchida, Yoshiharu Tasaka

Fumikazu Ikazaki, Motoo Yumura

*National Institute of Materials
and Chemical Research,
Tsukuba 305, Japan*

Atsuo Kasuya, Yuichiro Nishina

*Institute for Materials Research,
Tohoku University, Sendai 980-77, Japan*

1. Rinzler, A. G. *et al.* *Science* **269**, 1550–1553 (1995).
2. de Heer, W. A., Chatelain, A. & Ugarte, D. *Science* **270**, 1179–1180 (1995).
3. Collins, P. G. & Zettl, A. *Appl. Phys. Lett.* **69**, 1969–1971 (1996).
4. Gomer, R. *Field Emission and Field Ionization* Ch. 2 (Harvard Univ. Press, Cambridge, MA, 1961).
5. Ikazaki, F. *et al.* in *Proc. 211th ACS Natl Meeting* Vol. 42 (eds Cronauer, D. C. *et al.*) 113–117 (Am. Chem. Soc., Washington DC, 1996).

DNA fingerprinting from single cells

Van Oorschot and Jones reported in *Scientific Correspondence*¹ that short tandem repeat (STR) profiles (DNA fingerprints) can be obtained from cells left on pens, car keys, and so on. Although this technique is a dramatic breakthrough with important implications for forensic science, there are two main limitations. First, more than 1 ng DNA (equivalent to 200 cells) is required; and second, a single STR locus is used, which does not provide much information. Here we report a system for

determining STR profiles from single cells using six forensic STR markers, which we believe is the first time that single cells have been typed using modern forensic techniques.

Genetic profiles can be obtained from single cells using an STR profiling system already in routine use in the United Kingdom, which gives a matching probability of roughly 1 in 50 million. The STR profile is reliable and accurate and is obtained within 5–6 hours.

We analysed 226 buccal cells from four different individuals, isolating each cell using micromanipulation procedures. DNA was amplified using a routine forensic identification system with modifications (different primer concentrations, Ampli-Taq Gold, and 34 cycles)². This system uses fluorescent polymerase chain reaction (PCR) amplification of six STR markers to provide an STR profile, and amelogenin to diagnose sex. We compared single-cell results to known STR profiles from the cell samples. We amplified DNA in 91% (206/226) cells, obtaining a full DNA profile in 50% (114/226) and an acceptable profile (four or more STRs) in 64% of these cells (Table 1). The remaining 27% of cells gave incomplete, partial profiles. Although these partial profiles may be insufficient for typing, they may provide sufficient information to exclude potential suspects.

These results show great promise and could be applied, for example, to smudged fingerprints, single flakes of dandruff, single sperm in multiple rape cases, and small samples left on weapons or vehicles (see ref. 1). However, they must be considered with some caution. Although we obtained an acceptable profile in almost two-thirds of cells tested, care must be taken with the interpretation of the results. Stochastic effects cause preferential amplification and

Table 1 Details of analysis

Number of single cells analysed	226
Results obtained	206 (91%)
Amplification failure	20 (9%)
Full STR profile	114 (50%)
Acceptable profile (amelogenin, ≥ 4 STRs)	144 (64%)
Partial profile (1–4 STRs)	62 (27%)
Surplus alleles*	28 (12%)
False alleles**	11 (5%)
Allele dropout	88 (39%)

*Additional allele present in conjunction with true alleles.

**Additional allele in place of true allele. Extra-allelic peaks could be caused by contamination, somatic mutation or PCR-generated non-allelic peaks. We never saw more than two additional peaks in a profile or in 18 negatives, minimizing the possibility of cellular contamination. When surplus alleles were observed we considered the locus, but not the profile, uninformative. We observed allele dropout in 39% of cells at a rate of $\sim 10\%$ in each allele. If two cells are analysed then the risk of allelic dropout and misinterpretation in cells is reduced to 1%, if three cells 0.1%, and so on. Wild-card designations and conservative statistical criteria are needed to ensure that evidential value can be properly assessed.

allele dropout³. Consequently, this technique requires validation before use in forensic casework. We are currently investigating a robust interpretation strategy for single-cell STR profiling. This work also raises issues about STR profiling low numbers of cells and the need for stringent precautions in collecting and processing to avoid contamination.

I. Findlay, A. Taylor, P. Quirke

Department of Molecular Oncology,
Algernon Firth Building, University of Leeds,
Leeds LS2 9LN, UK

R. Frazier, A. Urquhart

Forensic Science Service, Gooch Street North,
Birmingham B5 6QQ, UK

1. van Oorschot, R. A. H. & Jones, M. K. *Nature* **387**, 767 (1997).
2. Sparkes, R. et al. *Int. J. Legal Med.* **109**, 195–204 (1996).
3. Findlay, I., Ray, P., Quirke, P., Rutherford, A. J. & Lilford, R. *Hum. Reprod.* **10**, 1609–1618 (1996).

Archaeopteryx-like skull in Enantiornithine bird

The bird *Cathayornis* from the Early Cretaceous period gives the first evidence for a post-Jurassic survival of an *Archaeopteryx*-like skull in birds. This skull combines short, toothed premaxillaries, nasals meeting at the midline and submaxillary fossae in the antorbital fenestra.

During the late Mesozoic era, from the Early Cretaceous to the latest Cretaceous (Maestrichtian), two distinct groups of birds co-existed as separate lineages^{1,2}. One of these, the ornithurine birds, survived into the Cenozoic era to give rise to all modern birds^{3,4}. The other lineage, the Enantiornithes, became extinct along with dinosaurs at the end of the Cretaceous^{3,5}.

The structure of Enantiornithine bird skulls is poorly understood. Until now, the best known was *Gobiapteryx*, of which there are fragmentary adult and embryonic skulls⁶ showing the absence of teeth, a reduced antorbital fenestra and an *Archaeopteryx*-like quadrate.

Cathayornis from the Early Cretaceous of China has the typical elongated outer metacarpal and the characteristic shapes of the scapula, coracoid, and distal tibiotarsus found in later enantiornithine birds^{3,7}. Photographs of the skull have been published⁸ and a drawing of it as it is preserved⁷, but there has been no detailed reconstruction of any enantiornithine skull published to date. Our reconstruction (Fig. 1) is based on the holotype of *Cathayornis* (specimen IVPP, V9769) and skulls from two skeletons referred to that genus (IVPP, V10896 and V10916). The latter two skulls provide information on the maxilla, lacrimal and a lateral view of the quadrate. The premaxilla bears four or five small teeth that are directed downwards or slightly forwards. It is covered externally with large nutrient foramina. The dorsal process of the premaxilla extends posteriorly slightly beyond the edge of the nares (nostrils). It is slightly shorter in *Archaeopteryx* (Fig. 1 c,d). The nares meet on the midline as in *Archaeopteryx* and are overlapped by the premaxillaries on their anterior process.

The premaxillaries are toothed and about one-third the length of the skull compared to about one-quarter in *Archaeopteryx*, and the nasals are shortened. The antorbital fenestra is large and triangular with two distinct anterior maxillary fossae. The maxillary is toothed. Teeth are set in sockets indicating that the individuals are mature. The teeth are not serrated and have

the expanded base and waisted crown typical of all known toothed birds (Fig. 1g). The 'T'-shaped lacrimal is inclined posteriorly. The braincase is expanded over that in *Archaeopteryx* (Fig. 1 b,d) and this may reflect an increased brain size coupled with an improved shoulder girdle (keeled sternum) and presumably more sophisticated powers of flight in *Cathayornis*. There is a bone in the posterior corner of the skull that might be a squamosal, but the quadrate articulation and basicranial region is not preserved in our material, nor is the quadratojugal and jugal.

There is a well-preserved quadrate (Fig. 1f) lying disarticulated and behind one skull (IVPP, V10916). It is a long slender bone with a single small proximal head and very little orbital process. It is similar to the quadrate (Fig. 1e) of the London, Eichstätt and seventh *Archaeopteryx* specimens⁹ and to that of *Gobiapteryx*. In *Archaeopteryx* the dorsal process of the jugal is too far back to contact a postorbital even if one were present¹⁰. The reduction or loss of the postorbital frees the jugal bar for prokinesis and may have occurred more than once in avian evolution. *Cathayornis* lacks evidence for a postorbital but one is present in *Confuciusornis*. The flattened nature of the nasal and the nasal process of the premaxilla⁷ may also indicate the presence of prokinesis in *Cathayornis*.

If *Archaeopteryx* and the Enantiornithes are united into a monophyletic Sauriurae¹, then the presence of a primitive *Archaeopteryx*-like skull in *Cathayornis*, after the derived postcranial differences between enantiornithine and ornithurine birds was established, indicates that the modern ornithurine skull and 'typical avian kinesis' was developed independently by ornithurine birds.

**Larry D. Martin
Zhonghe Zhou***

Natural History Museum
and Department of Systematics and Ecology,
University of Kansas, Lawrence,
Kansas 66045, USA

*and Institute of Vertebrate Paleontology and
Paleoanthropology,
Chinese Academy of Sciences, PO Box 643,
Beijing 100044, China
e-mail: ldmartin@falcon.cc.ukans.edu

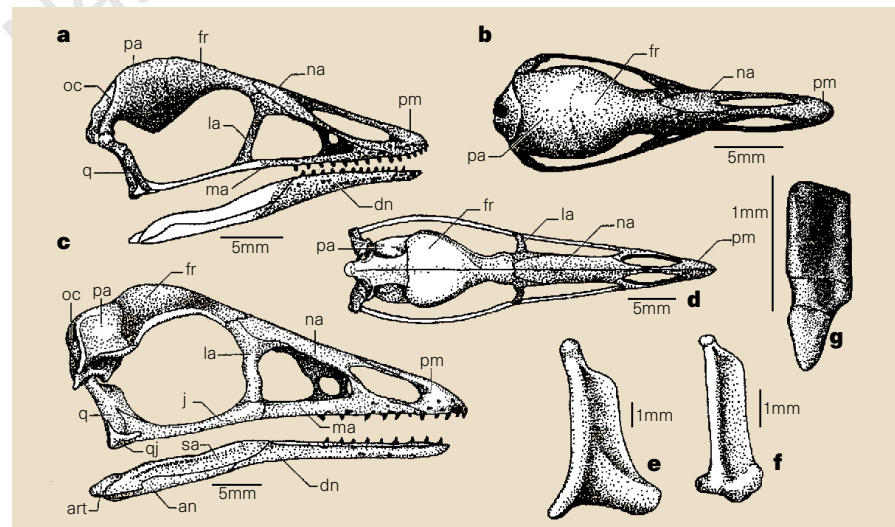


Figure 1 Reconstruction of skull structures of *Cathayornis* and *Archaeopteryx*. **a**, Lateral view of *Cathayornis*. **b**, Dorsal view of *Cathayornis*. **c**, Lateral view of *Archaeopteryx*. **d**, Dorsal view of *Archaeopteryx*. **e**, Quadrate of *Archaeopteryx*. **f**, Quadrate of *Cathayornis*. **g**, Labial view of tooth from the maxillary of *Cathayornis*. Abbreviations: an, angular; art, articular; dn, dentary; fr, frontal; j, jugal; la, lacrimal; ma, maxillary; na, nasal; oc, occipital; pa, parietal; pm, premaxillary; q, quadrate; qj, quadratojugal; sa, surangular.

1. Hou, L., Martin, L. D., Zhou, Z. & Feduccia, A. *Science* **274**, 1164–1167 (1996).
2. Martin, L. D. in *Origin of the Higher Groups of Tetrapods* (eds Schultze, H. P. & Trueb, L.) 485–540 (Comstock, Ithaca, 1991).
3. Martin, L. D. *Cour. Forschungsinst. Senckenberg* **181**, 23–36 (1995).
4. Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, New Haven, 1996).
5. Chiappe, L. M. *Nature* **378**, 349–355 (1995).
6. Elzanowski, A. *Paleontol. Polon.* **42**, 147–179 (1981).
7. Zhou, Z. *Cour. Forschungsinst. Senckenberg* **181**, 9–22 (1995).
8. Zhou, Z., Jin, F. & Zhang, J. *Chinese Sci. Bull.* **37**, 1365–1368 (1992).
9. Wellnhofer, P. *Archaeopteryx* **11**, 1–47 (1993).
10. Elzanowski, A. & Wellnhofer, P. *J. Vert. Paleontol.* **16**, 81–94 (1996).

Mental modules on the brain

How the Mind Works

by Steven Pinker

Norton: 1997. Pp. 636. \$29.95. To be published in the United Kingdom by Allen Lane in February

P.N. Johnson-Laird

How does your mind work? Here is a test devised by Peter Wason to help you to find out. I put four cards in front of you, bearing respectively: 'A', 'B', '2' and '3'. Each card has a number on one side and a letter on the other side. You have to choose which cards to turn over to determine whether the following rule is true or false about these cards: if a card has an 'A' on one side, then it has a '2' on the other side. Remember your choice, because the test is a linchpin for a central idea in Steven Pinker's new book.

Pinker is distinguished among psychologists for research on vision and imagery, and language and learnability. But he is famous for his last book *The Language Instinct* (Morrow/Viking, 1994), which was a brilliant success and bestseller. He has a gift for exposition, for witty analogies, for apposite quotation from his vast knowledge of culture, high and low. And he knows when to skip the details of difficult technicalities. He uses all these skills in his latest book. It is erudite, light in touch, but long.

What it defends is a threefold package of ideas. First, the mind works like a computer: both process information. Second, unlike existing computers, it is divided into separate modules specialized for different tasks, such as stereopsis and reasoning about social exchanges. Third, the high road to scientific understanding is to follow the 'evolutionary psychologists', and to reverse-engineer the way in which the mind adapted to the lives of our evolutionary ancestors.

Whether the mind is a computational device is an open question, but it is axiomatic among cognitive scientists that many of its functions can be modelled computationally, and that the development of such programs provokes both theorizing and empirical investigations. Pinker shows, for example, how the understanding of stereopsis was advanced by the late David Marr's insistence on thinking about what the mind has to compute, how it carries out the computations and how such an algorithm is embodied in neurons.

Evolution can hardly have shaped a mind that depended on one homogeneous computational glob. It must tinker with a large set of separate mental organs, and so the mind must be divided into functionally separate modules, for much the same reason that programmers divide their programs into separate routines. The modular claim goes

back to nineteenth-century neurology, but it has been challenged — most recently, by 'connectionists'. The success of their learning algorithms for homogeneous networks of toy neurons, they say, calls for psychologists to rethink innateness.

But 'raw connectoplasms', as Pinker replies, cannot cope with significant aspects of mental life. It fails, so far, with recursion, and accordingly with the way in which the meanings of sentences are composed from the meanings of their parts, which in turn are composed from the meanings of their parts, and so on, down to the meanings of morphemes. Yet connectionists still seek an algorithm that will learn natural language. A more realistic target would be one that learns a simple logical calculus.

The most controversial aspect of Pinker's package is his defence of evolutionary psychology. Like Darwin, most cognitive scientists believe that the brain is a result of evolution. Most of them, however, are bemused witnesses to the bare-knuckle fight now bloodying the pages of the *New York Review of Books*. In one corner are orthodox neo-Darwinians such as John Maynard Smith, and their seconds, the evolutionary psychologists, who believe one big thing: natural selection is the sole governor of evolution. In the other corner are those, such as Stephen Jay Gould, who reject the orthodoxy along with sociobi-

ology and its descendant by modification — evolutionary psychology. The pugilists should not be allowed to bite off this nub: you can reject evolutionary psychology without rejecting neo-Darwinism. "Selection no doubt accounts for certain details, but in all probability not for the general character of mind" — so wrote J. B. S. Haldane.

The evolutionary psychologists, Pinker explains, argue that mental modules are adaptations to the lives of our forebears. Thus, Leda Cosmides, a leader of the group, proposed an innate module for reasoning about cheating, because social exchanges were important to the hunter-gatherers of the Pleistocene epoch.

What did you select in the four-card test? Most people correctly choose the 'A' card, and some choose the harmless '2' card too. Few realize the need to choose the '3' card; yet if it has an 'A' on its other side, then the rule is false. Cosmides replaced the letters and numbers with a problem about a potential cheater, and then, as she predicted, most people made the correct choice.

There is an alternative explanation, however. Perhaps your mind normally reasons about what is true, not what is false. So, any manipulation that helps you to consider false instances of the rule, including reminders about cheating, should improve your performance. Half a dozen studies have corrobo-

Homage to Patagonia



The Selk'nam men of Patagonia disguised themselves as vengeful spirits to scare women from trying to seize power. In the event, of course, that was done by European colonists, who put a price on the heads of the native peoples, paying extra for the ears and fetuses of pregnant women. Ten thousand years ago the ancestors of the Selk'nam and other tribes left delicate hand outlines on cave walls. A gripping read and lavishly illustrated, *Patagonia: Natural History, Prehistory and Ethnography at the Uttermost End of the Earth* edited by Colin McEwan, Luis A. Borrero and Alfredo Prieto (British Museum Press, £14.99 (pbk)) accompanies a major exhibition at London's Museum of Mankind, which runs until 31 December.

rated this explanation without depending on a check for cheaters. Pinker does not describe these studies.

Which modules are adaptations? Stereopsis, no doubt; language, perhaps; music, probably not. As sceptics say, we cannot go back to the Stone Age to test natural selection at work. The best we can do is to tell a "Just So" story about what should have been adaptive for our ancestors, and to test whether it predicts how our minds work.

Pinker does indeed spin some entertaining stories — from how humans got their taste for floral landscapes to how siblings get into spats. Indeed, the range of his book is formidable, embracing society, art and religion.

But what do we do when a "Just So" story just isn't so? Allow, like Pinker, that not all mental functions are adaptations, or rewrite the story to fit the facts? Evolutionary psychologists could argue that truth was more important than falsity in the evolution of communication. Indeed, it probably was, but it does not follow that an innate module evolved for reasoning about the truth. The bias may be a result of experience or a by-product of some other adaptation. Pinker's evolutionary psychology is a useful heuristic for generating hypotheses, but not a refutable theory. □

P. N. Johnson-Laird is in the Department of Psychology, Princeton University, Green Hall, Princeton, New Jersey 08544, USA.

Old views of the past

The Ancient Kingdoms of Peru

by Nigel Davies

Penguin: 1997. Pp. 221. £8.99, \$13.95 (pbk)

Nicholas J. Saunders

Archaeology plays on our fascination with the past. Add geographical distance to that of time and arguably nowhere is as intriguing as South America. Making sense of the heady mix of science, history and romantic imagination is the role of popular books. In *The Ancient Kingdoms of Peru*, Nigel Davies responds by offering a competent, sometimes idiosyncratic overview, weighed down by an outdated approach focusing on a handful of apparently disarticulated civilizations.

The defining fault is a structure that straitjackets some 8,000 years and an astonishing diversity of prehistoric societies into the first hundred pages — barely half the book. The last hundred pages are devoted to the Inca, despite (perhaps because of) the author's recent book on this topic. So we have a rehearsed Inca book, with cursory accounts of pre-Inca cultures bolted onto the front. Previously, bias in knowledge might have justified this, but not today.

Given the small amount of space that the author gives himself, there is a surprising amount of repetition and marginal informa-



Secrets of Peru's past unearthed: Chimor gold ceremonial knife, eleventh to fifteenth century AD (left) and Nazca effigy jar of a deity, 300–600 AD.



tion. We have just three pages of text dealing with the precocious Chavín civilization (800–200 BC) and no mention of its shamanic, perhaps hallucinogenically inspired iconography and ideology which so influenced many subsequent civilizations. Compared with this meagre coverage, the account of the less important but more famous Nazca culture is five times longer. Davies's interest seems fuelled by the array of lines and animal figures that cover the Nazca desert. An inordinate amount of space is occupied in recounting (even analysing) spurious details of Von Däniken's extraterrestrial and other fringe ideas. Yet interspersed with such nonsense are useful vignettes of more serious investigations that see the desert markings as related to ritual activity, rain and fertility in time-honoured Andean tradition.

Unfortunately, Von Däniken reappears in the chapter on the great pre-Inca ceremonial centre of Tiahuanaco, built on the southern shores of Lake Titicaca. This is followed by a recounting of equally wacky theories about crashing planets and huge floods. Why Davies should choose to acknowledge, let alone revive, such ideas and embed them in an otherwise sober account of this major pan-Andean civilization is its own mystery.

More successful is the treatment of the broadly coeval Andean city of Huari, which flourished around AD 600. This comparatively little-understood polity occupies a crucial position in the prehistory of Andean civilization, and Davies picks his way through the often confusing literature to present a fair picture of Huari and associated sites such as Pikillacta. Similarly successful, although again too brief, is the account of the kingdom of Chimor on Peru's north coast, its great canals and mudbrick architecture, its

(usually militarily agonistic) relationships with other north-coast cultures, and the disruptive role of an El Niño climatic event in AD 1100.

The Chimor capital of Chan-Chan fell to the Incas around AD 1470, and, from this point, Davies launches into the blend of history and archaeology with which he evidently feels more at ease. The four Inca chapters that follow are generally informative and occasionally perceptive but, again, too much is missing. Failure to convey the excitement and breadth of modern investigations is everywhere apparent, not least in the selection of photographs and their minimalist captions. These add nothing to a text sporadically illustrated with poor-quality line drawings, several reproduced upside down.

The most serious criticism, however, must be the lack of any sustained feel for the dramatic technological and theoretical advances that have redefined archaeology over the past 30 years. Art history, ethnographic analogy, scientific analysis and sophisticated interpretive frameworks exist hardly at all in this book. Where, one might ask, are discussions on uniquely Andean concepts of symbolic landscapes, the ideology of gender and ancestor worship, cultural astronomy, household archaeology, the shamanic significance and technologies of metalwork, or the use of textiles to define and represent status, power and cosmology? Such subjects are inherently exciting and topical for a general audience yet are virtually ignored here.

In some ways this is not a bad book — it covers the ground, mentions obvious new work and gives a taste of Peru's stunning archaeological heritage. Yet it is also a lost opportunity. As Penguin's successor to John

Alden Mason's 30-year-old and endlessly reprinted classic, *Ancient Civilizations of Peru*, this is embarrassingly lightweight, with all the hallmarks of a rush job.

Ironically, Davies is an internationally respected authority on the Aztec and Toltec civilizations of Mexico, on which he has published many scholarly and erudite general accounts, informed by a lifetime's study of that material. South America, however, is not Mesoamerica, and this book offers little other than a glancing skim across one of the world's greatest archaeological regions. □

Nicholas J. Saunders is in the Department of Anthropology, University College London, Gower Street, London WC1E 6BT and the Department of Archaeology, University of Southampton, Southampton SO9 5NH, UK.

Wing commanders

Fly Pushing: The Theory and Practice of *Drosophila* Genetics

by Ralph J. Greenspan
Cold Spring Harbor Laboratory Press: 1997.
Pp. 155. \$35

***Drosophila* Cells in Culture**

by Guy Echalié
Academic: 1997. Pp. 702. \$135, £105

Ernst Hafen

The evolutionary conservation of biological processes discovered by developmental geneticists has promoted a multidisciplinary approach to biological problems. For example, studies on intracellular signal transduction are no longer carried out exclusively in the tissue culture dish but now also use complementary genetic approaches in model organisms such as the fruitfly *Drosophila*. Consequently, in the far corners of many large biochemistry laboratories, stations for pushing flies often compete for space with fraction collectors and DNA sequencers.

The renewed excitement in fly genetics calls for a new generation of scientists who feel at home working with these two different experimental systems. But just how does a graduate student with a PhD in biochemistry go about learning genetics in a fly lab? How can he or she write a sensible fellowship application without knowing the basic techniques and terminology used by fly geneticists? Standard books on *Drosophila* genetics and development are multivolume series of a thousand pages or more. Although Peter Lawrence's *The Making of a Fly* (Blackwell Scientific, 1992) provides a good introduction to the concepts of developmental genetics, it does not describe how to perform a cross or how and why stocks should be made genetically uniform before starting mutagenesis.

Fly Pushing fills this gap. Ralph J. Greenspan has written a concise summary

of the genetic terminology and techniques of *Drosophila* research. Easily read in an afternoon, this slender volume will help the novice to overcome initial difficulties in fly genetics. It starts with a description of the fly chromosomes, the most commonly used markers and how to handle flies. The chapter on mutagenesis discusses the advantages and disadvantages of different mutagens and gives examples of how to isolate new alleles of a given gene. The different approaches are described with equal weight. It would have been helpful here to add a section on how to identify mutations in a gene encoding the fly homologue of a biochemically well-characterized protein whose function is nevertheless unknown.

An important and central part of the book is the chapter on the synthesis of specific genotypes. Making specific genotypes is a daily problem in fly genetics; different mutations have to be combined in the same fly strain to test for interactions between genes. The construction of these stocks can easily take more than five generations. So it is important that the crossing scheme is right and that the crosses involve enough flies to obtain the desired end progeny. The book gives examples and practical advice on this topic.

The last two chapters deal with the analysis of mutant phenotypes. After a useful section on the terminology and classification of different types of mutations, methods of analysing mutant phenotypes at different stages in development are briefly discussed. It is obvious from the merciful brevity of the book that this discussion is by no means comprehensive. The author then describes different methods for creating genetic mosaics. For my taste, there is too much emphasis here on classical approaches such as gynadromorph mapping and segmental aneuploidy and too little emphasis on the more sophisticated techniques of modern genetic manipulation such as P-element transformation and



The biologist's best friend.

the GAL4 and FLP-FRT systems.

With problems presented after each section, this book provides a hands-on introduction to fly genetics. It is a must for all postdocs or graduate students moving into fly genetics and should be in the library of every *Drosophila* lab. I only wish there was something similar for fly people moving into biochemistry or cell biology.

Fly Pushing condenses the essentials of 75 years of fly genetics into 155 pages. By contrast, *Drosophila Cells in Culture* by Guy Echalié — the grand master in the field — includes in 702 pages everything known about *Drosophila* cell culture. This fact alone shows how much the field of *Drosophila* cell culture has been neglected over the past 30 or 40 years. The widely held attitude in the fly community of "why do it in cell culture if one can do it in the real animal" has kept people from developing this area of research. From the highs of *Drosophila* cell-culture research in the 1960s and 1970s, the same cell line (the S2 Schneider cell line established by Imogene Schneider) has survived to today, and has recently been rediscovered as a transfection vessel also for constructs encoding vertebrate proteins. In contrast to the widely used vertebrate cell lines, *Drosophila* S2 cells lack the corresponding vertebrate protein and so provide a null mutant background.

With few exceptions, no new cell-line culture techniques have been developed for *Drosophila*. This has created a void in cell-culture lines and knowhow for establishing and maintaining new cultures. Fly workers therefore look with great jealousy at the advances in the cell-culture systems in vertebrates. For example, the availability of a totipotent embryonic stem cell line in *Drosophila* would be essential for establishing homologous recombination in flies.

The first part of Echalié's book provides a comprehensive overview of the culture conditions of *Drosophila* cell lines, the methods for establishing new cell lines, and the available cell lines and their characteristics. The second part describes how cell-culture systems have been important in defining the biochemical properties of proteins such as adhesion molecules or cell surface receptors whose function has been inferred solely from the mutant phenotype. The chapter on the introduction of foreign DNA into cultured cells is of great practical value because it discusses and evaluates the different transfection methods and the use of different selection procedures and different promoters. The chapter ends with a series of detailed protocols for the various transfection methods.

Beginners with basic training in handling cultured cells will find the book a useful introduction to the art of culturing *Drosophila* cells. It is also the definitive

reference work on *Drosophila* cell culture. Cross-references to vertebrate cell culture systems and techniques put the topic in a wider perspective. May the book inspire many scientists to establish new cell lines, and so help to bridge the gap between vertebrate biochemistry and *Drosophila* genetics. □

Ernst Hafen is at the Zoologisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

Doing the right thing?

Research Ethics: A Reader

edited by Deni Elliott and Judy E. Stern
University of New England Press: 1997.
Pp. 336. \$25, £18.95 (pbk)

The Ethics of Scientific Research: A Guidebook for Course Development

by Judi E. Stern and Deni Elliott
University of New England Press: 1997.
Pp. 128. \$15, £10.95 (pbk)

John Galloway

There is an interesting tension between the romantic view of research as a disinterested quest for truth and the anything-but-disinterested motives of those actually doing the research. Science is as good a way to fame and occasionally fortune as any other sphere of human activity, and just as vulnerable to human frailty: greed, vanity, envy, jealousy and corruption. But does science have enough distinct problems to warrant its own moral and ethical code to guide — or control — the behaviour of its practitioners? After all, moral and ethical codes are defining features of other recognized professions such as medicine and law. Cynical observers will recognize the formulation of such codes as key 'public relations' steps on the ladder to professional status — and power.

Science's moral issues provide the theme of *Research Ethics*. This collection of essays and articles is accompanied by *The Ethics of Scientific Research*, a manual for setting up a teaching course in the subject. They cover the ground, although, inevitably given the format, unevenly; and both suffer from a general lack of cohesion.

Science poses fundamental ethical problems that the books can only attempt to address because these problems have no definitive solution. The human condition rules it out. Science continually subverts the moral codes by which we all live, dictates how we live, and how and when we die. The pace of scientific progress outstrips society's ability to cope with it. Scientists have to acknowledge that they have the potential to create evil as well as good — and that all progress must be paid for.

Recently, the media have been only too happy to add 'sin in science' to the long list of the world's other sins that they feel the public

need to know about. Institutions have also responded to the realization that there may be serpents in the scientific Garden of Eden. The penalty now, as then, for those caught transgressing is being cast out.

Guidelines on scientific conduct and the way to deal with suspected misconduct proliferate. Scientists face the same problems in dealing with dishonesty as society does in general. Misconduct needs to be defined and not just left for experts to recognize; there is also the small problem of proving fraud once it is suspected. Even more difficult is knowing how to deal even-handedly with those involved, particularly when the stakes are high. Inevitably there will be a tendency for scientists to see little wrong when their behaviour profits them personally. Nor is science separated from the rest of the world by some moral cordon sanitaire. It overlaps with, and is part of, other social systems — business, national security, politics — whose codes of behaviour are removed from those of an idealistic scholar pursuing truth for truth's sake. And science has its own politics — unavoidable once large sums of money are competed for, spent or made — with its own combination of self-interest and self-delusion.

Research Ethics brings out all these issues, although they are not always dealt with entirely objectively. Some of the articles seem too personal and self-justifying, as in the attack on the US Office of Scientific Integrity, apparently for not prosecuting cases impartially. Whether or not this criticism is justified, such an attack is surely out of place in a book such as this.

Scientific research creates new knowledge about the world — 'data' as scientists call it. And the demon that drives scientists is the challenge of being the first to make a particular discovery. As Peter Medawar pointed out, no one gets any credit for coming second. It is not surprising then that misconduct involving data is the equivalent of being caught horse-stealing in the Wild West, with similar consequences. Your career may be 'lynched', even if you are not. Most cases of scientific fraud involve scientists who are too creative with their data, taking liberties with their findings, helping them along if they do not support their pet theory — and occasionally just making them up. Guidelines increasingly focus on the need for scientists and laboratories to keep accurate records about their work and methods. *Research Ethics* describes a number of recent instances.

For those embarking on a career in research, their first experience of a moral issue may be one of ownership — whose research is it? Are the research findings you seem to have made yours, or do they belong to your supervisor or even the director of the laboratory? There are scientists who insist that their name appears on everything published from their laboratory, irrespective of

their input. Occasionally this behaviour results in the perpetrator being hoist with his own petard when his name appears on a paper later proved to be fraudulent. The patron saint of postdocs probably smiles broadly. There is an issue here, and one that science is starting to address even though the issue is not entirely straightforward. Perhaps the nearest counterpart to laboratories are the workshops of Renaissance painters. As the art critic Brian Sewell comments about the artist Lucas Cranach, "he commanded a considerable workshop... and his hand must have been rare in the many variants and replicas that were the staples of his business". And as the head of the Oxford laboratories where I worked said, "Be realistic, the only thing that matters is what I say about you and your work". He was probably right.

It is not obvious that dishonesty in science is inherently more serious than in any other walk of life, although clearly it manifests itself in particular ways. It is also worth remembering that if the results of research are important, any deception will invariably come to light. Dishonesty in science is therefore probably more stupid than wicked when one considers the price paid for being caught. But this is no excuse for science not keeping its own house in order. Stupidity can cause as much damage as deliberate dishonesty. That said, it is important to keep a sense of proportion. A clear message from these books is how easily things can get out of hand if procedures designed to deal with these problems are not tempered with common sense. □

John Galloway is at the Eastman Dental Hospital, 256 Gray's Inn Road, London WC1X 8LD, UK.

New in paperback

Rain of Iron and Ice: The Very Real Threat of Comet and Asteroid Bombardment

by John S. Lewis
Addison-Wesley, \$13

"His writing is agreeable and his scholarship impressive", S. Battersby in *Nature* 381, 124 (1996).

The Three Big Bangs: Comet Crashes, Exploding Stars, and the Creation of the Universe

by Philip M. Dauter and Richard A. Muller
Addison-Wesley, \$13

Reviewed by S. Battersby in *Nature* 381, 124 (1996).

Rogue Asteroids and Doomsday Comets: The Search for the Million Megaton Menace that Threatens Life on Earth

by Duncan Steel
Wiley, \$16.95

"Engaging and worthwhile", M. A'Hearn in *Nature* 378, 106 (1995).

Molecular motors: structural adaptations to cellular functions

Joe Howard

Molecular motors are protein machines whose directed movement along cytoskeletal filaments is driven by ATP hydrolysis. Eukaryotic cells contain motors that help to transport organelles to their correct cellular locations and to establish and alter cellular morphology during cell locomotion and division. The best-studied motors, myosin from skeletal muscle and conventional kinesin from brain, are remarkably similar in structure, yet have very different functions. These differences can be understood in terms of the 'duty ratio', the fraction of the time that a motor is attached to its filament. Differences in duty ratio can explain the diversity of structures, speeds and oligomerization states of members of the large kinesin, myosin and dynein families of motors.

There is no let-up in the pace at which new molecular motors are being discovered. In the past ten years, the founding members of the three families of motor proteins—myosin from skeletal muscle, axonemal dynein from sperm, and conventional kinesin from brain—have been joined by hundreds of new-found relatives whose amino-acid sequences contain regions of high similarity to the force-generating 'motor domains' of the founding proteins (see reviews for kinesin^{1,2}, myosin^{3,4}, and dynein⁵). The extent of molecular diversity can be appreciated on realizing that the several isoforms of myosin II that are found in skeletal, smooth and cardiac muscle, as well as in virtually all non-muscle cells, all fall within one subfamily—and yet there are at least ten other myosin subfamilies. Many of these motor-related proteins have been shown to move *in vitro* (they really are motors) and 'knockout' experiments (in which expression of a motor protein is prevented by targeted disruption of the gene encoding it) have revealed a remarkable diversity of cellular functions^{1–5} (Table 1). Some members in each family function in large macromolecular assemblies (myosin II in muscle, axonemal dynein in sperm, and cytoplasmic dynein and various kinesin-related proteins in the mitotic spindle), whereas others function independently or in small numbers to move membrane-bound vesicles and organelles (conventional kinesin, kinesin II, cytoplasmic dynein and myosin V). In addition to being motors, some members of the myosin I, III and IX subfamilies may also be signalling molecules^{6,7}. Mutations in myosins cause human diseases

such as familial cardiac hypertrophy⁴ and Usher's syndrome, which is associated with vision and hearing loss³.

Meanwhile, the understanding of the molecular mechanisms of force generation by motor proteins has proceeded rapidly with the solution of the atomic structures of skeletal muscle myosin⁸ and conventional kinesin⁹, and with the recent development of optical trapping, atomic-force microscopic and fluorescence techniques, which allow the visualization and manipulation of single motor molecules as they move along filaments in the presence of ATP^{10–13}.

Now is a good time to take stock and ask what we have learned so far about the molecular mechanisms of force generation. Can we find common principles in the workings of muscle myosin and conventional kinesin that will help us understand the function and mechanisms of the other proteins within the motor families? I will argue that the new results have forced us to revise some of our fundamental assumptions about how motors, especially myosin, work, but they have also given us new insight into the divergent structures and cellular roles of motor proteins in general.

Conventional kinesin and muscle myosin

A superficial comparison of muscle myosin and conventional kinesin presents a puzzle: the structures are so similar yet the motor functions and mechanisms appear to be so different. Electron microscopy reveals two motors with a common domain organization: each has two identical globular 'heads', which dimerize

Table 1 Motor speeds *in vivo* and *in vitro*

Motor	Speed* <i>in vivo</i> μm s ⁻¹	Speed† <i>in vitro</i> μm s ⁻¹	ATPase‡	Function
Myosins				
Myosin IB	?	0.2	6	Acanthamoeba motility ^{71,72}
Myosin II	6	8	20	Rabbit psoas muscle ⁷³
Myosin II	0.2	0.25	1.2	Avian smooth muscle ^{74,75}
Myosin V	0.2	0.35	5	Vesicle transport in yeast and mice ^{66,67}
unknown	60	60	?	<i>Nitella</i> cytoplasmic streaming ⁷⁶
Dyneins				
Axonemal	7	4.5	10	Sea urchin sperm ^{77,78}
Cytoplasmic	1.1	1.25	2	Retrograde axonal transport ^{64,79,80}
Kinesins				
Conventional	1.8	0.8	44	Anterograde axonal transport ^{12,27,79}
Fla10/Kin11	2.0	0.4	?	Anterograde transport in flagella ^{81,82}
Kip1/Eg5	0.018	0.06	2	Mitosis and meiosis ^{83–85}
Ncd	?	0.09	1	Meiosis and mitosis ^{86,87}

**In vivo* speed applies to the motion of the motor relative to the filament without external load.

†*In vitro* speed is that of purified motor and filament at high ATP concentration.

‡ATPase is the maximum filament-activated rate of hydrolysis per head per second measured in solution at high ATP and filament concentrations.

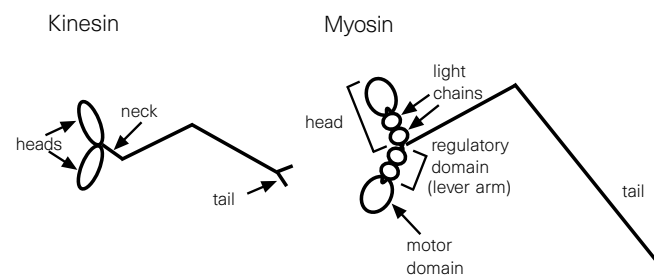


Figure 1 Structures of conventional kinesin and muscle myosin based on electron microscopy and sequence analysis. The kinesin homodimer has two heads, each about 8 nm long, which are joined at the coiled-coil neck, a dimerization domain. The neck connects to a long coiled coil that terminates in the tail region, thought to bind to the organelle cargo. Myosin's heads are about twice the size of kinesin's and are composed of two major domains, the motor domain which binds actin and nucleotides, and the regulatory domain, an 8-nm-long α -helix that is stabilized by two calmodulin-like light chains. The regulatory domain, which is thought to act as a lever arm, then connects to a long coiled-coil rod, which in muscle oligomerizes to form the thick filament.

through long coiled-coil dimerization regions that terminate in a 'tail' (Fig. 1). The heads are the crossbridges between the thin and the thick filaments of muscle¹⁴ and between organelles and microtubules in neurons¹⁵ and they contain the motor domains that bind both the nucleotide and the filament. Except for the dozen or so amino acids that bind the nucleotide, the head domains of myosin and kinesin contain no amino-acid sequence similarity. Thus it came as a surprise that the core 180 amino acids within the myosin and kinesin head structures have the same fold⁹. Furthermore, the preserved sequence direction and order of the α -helices and β -sheets within the core suggest that these motors evolved from a common ancestor.

Yet despite these structural similarities, there are profound functional differences between kinesin and myosin. Conventional kinesin operates alone or in small numbers to transport membrane-bound organelles over large distances (up to a millimetre) along microtubules¹⁶, whereas muscle myosin operates in huge arrays of up to a billion molecules in a large muscle fibre and moves relatively short distances (up to about one micron) along actin filaments¹⁴.

These functional differences are inherent to the motor proteins themselves, rather than being due to different accessory proteins in the different tissues. Surfaces sparsely coated with purified kinesin support the gliding of microtubules, and quantitative dilution studies show that a single kinesin molecule is sufficient for motility: one motor can walk up to several microns along the surface of a microtubule without falling off¹⁷. This property is an adaptation for kinesin's role as an organelle motor, and kinesin is said to be a 'processive' enzyme, by analogy to the DNA and RNA polymerases which move along and polymerize up to thousands of bases during each encounter with the DNA strand. By contrast, myosin assays do not work at low density, and it is estimated that a minimum of tens to hundreds of myosin heads are needed for the continuous movement of an actin filament in a gliding assay^{18,19}. Although any propensity of an organelle motor for dissociating from its filament would pose a transport problem, it is not a problem in muscle, in which there are 300 myosin molecules per thick filament²⁰ and where the sarcomere is reinforced to ensure that the muscle does not fall apart even when it is relaxed and none of the crossbridges is engaged.

High-resolution single-molecule experiments confirm that kinesin is processive whereas myosin is not. Using optical tweezers^{10,21} or glass fibres¹² to impose loads and photodiode detectors to measure displacements with nanometre sensitivity, it has been shown that a single molecule of kinesin can move hundreds of nanometres along a microtubule even against an opposing force of up to 5 piconewtons. By contrast, a single myosin molecule makes only transitory interactions with the actin filament and is unable to progress to the next binding site even when unloaded^{11,22,23}.

The duty ratio concept

The functional differences between kinesin and myosin can be understood using the concept of the 'duty ratio', the fraction of

the time that a motor domain spends attached to its filament. The concept first arose (initially called the duty cycle²⁴) to explain the high force generated by smooth muscle, an issue to which I will return. In order to move along a filament through a distance that is large compared to molecular dimensions, each crossbridge must cycle repetitively between attached and detached phases (Fig. 2a). In the attached phase, of duration τ_{on} , the head or crossbridge undergoes the 'working' stroke, and in the detached phase, of duration τ_{off} , it undergoes the 'recovery' stroke, which returns the crossbridge to its initial conformation (Fig. 2b). We define the working distance (δ) as the distance that the distal, cargo-binding end of the crossbridge moves relative to the proximal filament-binding end (Fig. 2b). By recovering during the detached phase, the motor avoids stepping back, and so progresses through the working distance during each cycle. We formally define the duty ratio, r , as the fraction of the time that each head spends in its attached phase:

$$r = \frac{\tau_{on}}{\tau_{on} + \tau_{off}} = \frac{\tau_{on}}{\tau_{total}} \quad (1)$$

Unlike a two-stroke internal combustion engine, which has a duty ratio of 0.5 because the working stroke (the expansion stroke) and the recovery stroke (the compression stroke) are constrained to be of equal duration, the duty ratio of a crossbridge can in principle be large (~ 1) if it spends most of its time attached, or small (~ 0), if it spends most of its time detached. Because a two-headed molecule of conventional kinesin is able to maintain continuous attachment to the microtubule, its duty ratio must be at least 0.5, otherwise there will be times when neither head is attached and the motor will diffuse away from the filament. On the other hand, because skeletal muscle myosin must be in large assemblies with at least 10 to 100 crossbridges to move, its duty ratio must be small, ~ 0.01 to 0.1, the reciprocal of the minimum number of heads needed for continuous motility^{19,25}.

In addition to explaining why some motors can operate alone but others must work in assemblies, the duty ratio also provides the crucial link between the chemical speed of a motor (its ATPase rate) and the mechanical speed (its velocity²⁶). Consider a filament moving at constant speed, v , over an array of fixed motor proteins, as occurs during filament sliding in muscle or in an *in vitro* gliding assay. We suppose that there are enough heads interacting with the filament to ensure continuous motility: that is, at least one kinesin molecule or at least 10 to 100 myosin molecules. If each head is attached for time τ_{on} and moves through the working distance δ , then $v = \delta/\tau_{on}$. On the other hand, because the cycle is driven by ATP hydrolysis, we expect that the total cycle time $t_{total} = 1/V$, where V is the rate at which each head hydrolyses ATP. Substituting these expressions for τ_{on} and t_{total} into equation (1), we obtain another expression for the duty ratio:

$$r = \frac{\delta \cdot V}{v} \quad (2)$$

It is important to note that this argument assumes that exactly one

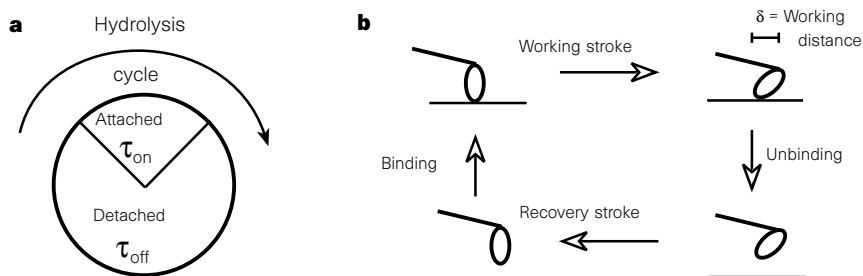


Figure 2 The crossbridge cycle. **a**, It is proposed that a crossbridge, or 'head', makes a working stroke during the attached phase (of duration τ_{on}) and makes a recovery stroke during the detached phase (of duration τ_{off}). **b**, By recovering its initial conformation while detached, the motor avoids stepping backwards and so progresses by a distance equal to the working stroke during each cycle. We define the duty ratio as the fraction of the time that the motor is bound and in its working phase.

ATP is hydrolysed per mechanical cycle: although there is indirect evidence that this is true for kinesin at low load^{27–29}, a pressing goal of the motor field is the direct measurement of the coupling stoichiometry at high and low loads.

The value of this analysis is that it explains some very puzzling kinetic differences between myosin and kinesin as simply being due to differences in duty ratios. The puzzle is that the unloaded gliding speed of fast skeletal muscle myosin is ten times greater than that of conventional kinesin, even though myosin's ATPase rate is only one half that of kinesin's (Table 1). In other words, myosin moves some twenty times farther than kinesin for each ATP that it hydrolyses. This leads to a paradox that generated much controversy in the myosin field^{30,31}: during the 50 ms it takes for each myosin head to hydrolyse a molecule of ATP, the actin filament moves through a distance of ~ 400 nm ($8,000 \text{ nm s}^{-1} \times 50 \text{ ms}$), a distance much larger than the dimension of the myosin head itself. The resolution of the paradox is that myosin has a low duty ratio: with $r = 0.01$ and an ATPase rate of 20 per head per second, a speed of $8,000 \text{ nm s}^{-1}$ can be reached with a working distance of only 4 nm (equation (2)), well within the dimension of the crossbridge. The crucial point is that each of the hundred or so heads moving the filament contributes only 4 of the 400 nm moved while it hydrolyses one molecule of ATP; then, while this motor is detached, the other heads sweep the filament along the rest of the way. By contrast, because kinesin has a high duty ratio, it needs a high ATPase rate to attain even moderate speeds: a working distance of 8 nm (see below), a speed of 800 nm s^{-1} , and an ATPase rate of 50 s^{-1} implies that $r = 0.5$, consistent with the duty ratio needed to account for kinesin's processivity.

Structural basis for the duty ratio

The definition of the duty ratio in terms of attached and detached times (equation (1)) gives the misleading impression that the duty ratio can equal any number between zero and one, depending on the biochemical 'on' and 'off' rate constants. However, this is not the case. In fact, the duty ratio can take on only a discrete set of values owing to the discrete sites on the filament to which the crossbridges can bind. To see why this is so, let d be the minimum distance between the motor's consecutive binding sites on the filament, the 'stepping stones'. We call this distance the path repeat distance, or simply the path distance: it depends not only on the structure of the filament, but also on the path that a particular motor follows on the surface of the filament. Now again consider a filament gliding over a fixed array of motor proteins. If the working distance, δ , is smaller than the path distance, d , then it is clear that each individual crossbridge must spend a significant time detached while other (attached) crossbridges move the filament and the next binding site forward. In this case, the duty ratio must be less than one, and continuous motility will require an assembly of crossbridges. In other words, the duty ratio ought to equal the fraction of the distance to the next binding site that the working stroke takes the crossbridge:

$$r = \frac{\delta}{n \cdot d} \quad (3)$$

where the integer n can be greater than one if the crossbridge skips over one or more of the stepping stones. Equation (3) represents the crucial steric constraint that is present in a moving, motor-filament system, but absent when isolated heads are freely diffusing in solution (in which case the path has no meaning). I will argue that this equation, which is both new and likely to generate some controversy, is of great use in understanding the diverse structures and biological functions of motors.

The first question to ask is whether this structural definition of the duty ratio accords with the previous definitions in terms of fractional binding and ATPase-to-speed ratios. The answer comes from recent single-molecule experiments which have elucidated the

paths trodden by conventional kinesin and muscle myosin, as well as the working distances. Most cytoplasmic microtubules are composed of 13 protofilaments that run parallel to the axis of the microtubule: these microtubules do not rotate when they glide over a kinesin-coated surface, whereas specially synthesized 12- and 14-protofilament microtubules whose protofilaments supertwist about the microtubule axis³² do rotate with the pitch and handedness of the supertwist³³. Thus kinesin's path is parallel to the protofilaments, and, because there is just one kinesin binding site per tubulin dimer^{34,35}, the path distance must be a multiple of 8 nm, the length of the dimers that form the protofilament (Fig. 3a). High-resolution tracking experiments of single (dimeric) kinesin molecules show that its steps are probably equal to $8 \text{ nm}^{10,28,36}$, although alternating 7-nm and 9-nm steps, which would be expected if kinesin were to 'waddle' along adjacent protofilaments, have not been ruled out. These data are not quite complete enough to derive a definitive value of the duty ratio, because there is still no direct measure of kinesin's working distance. However, they are quite consistent with a duty ratio of 0.5 if we take the working distance to be 8 nm (the step), the path distance to be 8 nm (the tubulin dimer length), and $n = 2$ in equation (3), because the second head must jump over the tubulin dimer to which the other head is bound (Fig. 3a).

The path trodden by myosin is also parallel to the axis of the actin filament (Fig. 3b). Actin filaments rotate little³⁷ or not at all³⁸ as they glide over surfaces coated with myosin: this shows that myosin does not move along the two-stranded helix as this would cause one rotation per 72 nm, the pitch of this helix. Now it is possible that the crossbridge is so flexible that it binds equally well to almost any actin subunit around the circumference of the filament. Alternatively, if the crossbridge has limited flexibility then it will bind only to correctly oriented subunits, and the path distance will be equal to 36 nm, the half-pitch of the helix (Fig. 3b). The latter, 36-nm path distance is directly supported by high-resolution single molecule experiments in which an actin filament is suspended between two beads that are each held in an optical trap and in which the filament is 'bowed' past a fixed myosin crossbridge: binding is only observed at multiples of about 35 nm³⁹.

Although the finding that kinesin moves parallel to the axis of the microtubule is not surprising (for example, this is the shortest path from one end to the other), the conclusion that myosin follows a path parallel to the actin filament is. Some earlier models assumed a parallel path⁴⁰, but it has more commonly been considered, especially in model diagrams, that myosin follows the two-stranded helix of actin^{41,42}, presumably because the consecutive binding sites would be considerably closer (only ~ 5.5 nm). However, the new conclusions, based on *in vitro* data, do accord with other biological and structural considerations. If a motor treads upon an angled path, a torque will be generated³³: this will disrupt the arrays of the filaments found in muscle or sperm, or it will cause transported organelles to spiral about their filaments, creating serious steric problems, especially if there are two filaments within an organelle-diameter of each other, as is usually the case in the cytoplasm.

The working distance of a single skeletal muscle crossbridge has been directly measured using high-resolution single-molecule techniques. In the absence of an external load, it is 4 to 6 nm (refs 22, 43). There has, however, been considerable controversy about the interpretation of these experiments, and some groups have inferred much larger working distances in the range of 10 to 20 nm (refs 11, 23, 44), although these high values are likely to have been inflated by the brownian motion of the actin filaments held in the soft optical traps⁴⁵. The smaller value is consistent with mechanical measurements in muscle showing that the tension in an active muscle is reduced to zero by a rapid shortening of the muscle by only 4–6 nm per half sarcomere^{46,47}. Considering that 2–3 nm is accounted for by the shortening of the actin filaments^{48,49}, the remaining 2–3 nm then corresponds to the average strain in the attached crossbridges (which will range from zero to the full 4–6 nm).

The small working distance of 4–6 nm compared to the large path distance of 36 nm is the structural reason why the duty ratio of skeletal muscle myosin is so small. The duty ratio can be no larger than about 0.1 to 0.2, and could be much less if myosin skips over neighbouring binding sites. Thus the structural data predict a low duty ratio, in accordance with the motility and biochemical data.

A low duty ratio of only ~10–20% for skeletal muscle myosin was originally deduced by electron paramagnetic resonance spectroscopy, which showed that only about 20 per cent of the heads in an active muscle were ordered, as would be expected if they were bound to the actin filament⁵⁰. But this conclusion met a lot of resistance from physiologists. The problem was that the stiffness of an activated muscle that is prevented from shortening (an 'isometrically contracting' muscle), or that is shortening only very slowly, is about 80% of that of a rigor muscle, in which most of the heads are attached⁵¹. If it is assumed that the stiffness is proportional to the number of attached heads, then the duty ratio must be close to one. However, the discovery that the actin filaments contribute about

half the compliance of muscle^{48,49,51,52} means that the stiffness is not proportional to the number of attached heads; a low duty ratio is therefore not inconsistent with the stiffness measurements, especially if the rigor heads are themselves somewhat less rigid than the active heads. On the other hand, a high duty ratio in muscle leads to a mechanical inconsistency: if all the crossbridges are bound and contributing to the stiffness of muscle, then the individual crossbridges cannot be very stiff and so will perform only a small amount of work per mechanical stroke⁴⁷. This means that each crossbridge must make many working strokes for each ATP that it hydrolyses as muscle has an efficiency²⁰ of 50%. However, multiple steps are not observed in the single-molecule recordings. The discrepancy is resolved if the duty ratio is low⁵³, because in this case the individual crossbridges are nearly an order of magnitude stiffer⁴² and so generate more force and work. The higher stiffness of the crossbridges also reinforces the view that they lack the flexibility to bind anywhere along the two-stranded actin helix. Finally, there is direct evidence that the low duty ratio in muscle is due to steric rather than

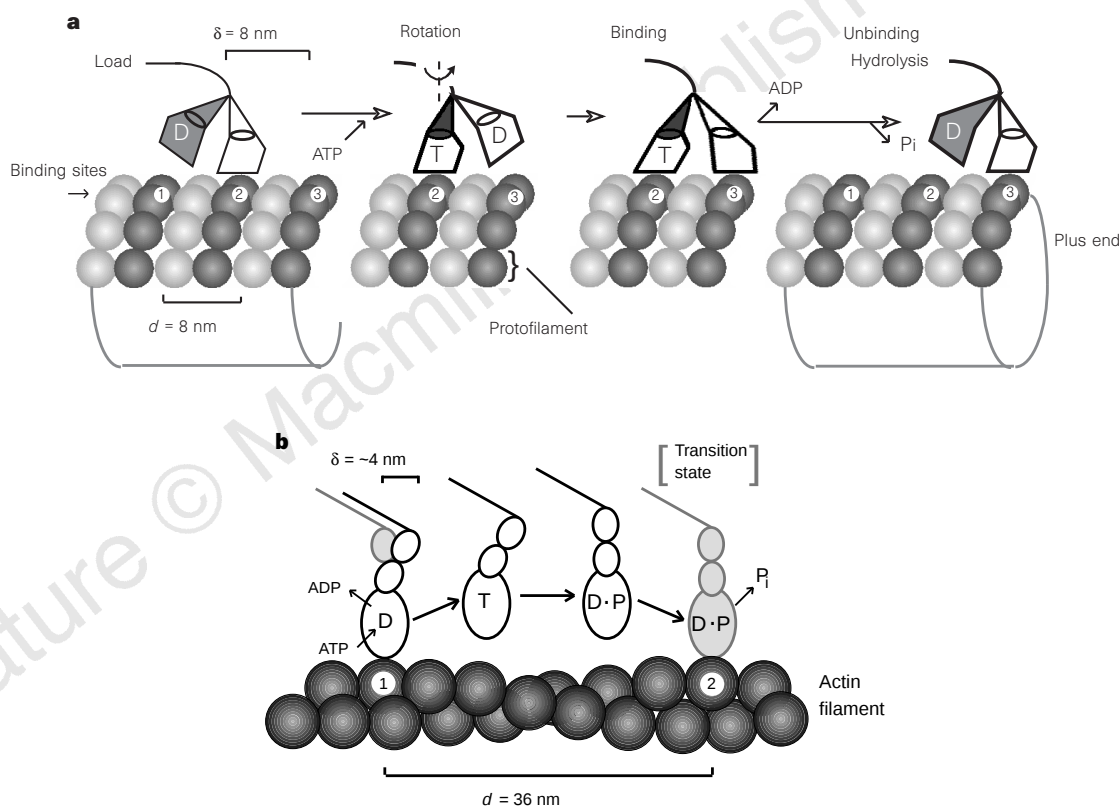


Figure 3 Structural and chemical models for the movement of **a**, kinesin, and **b**, myosin. **a**, The motion of conventional kinesin along the microtubule surface. Kinesin's path follows one of the protofilaments. As there is one binding site per tubulin dimer (α - and β -subunits shown as light and dark spheres, respectively), the path distance, d , is 8 nm, the dimer's length. The working distance is also 8 nm, the length of the step made by the distal, cargo-binding end of the motor during the hydrolysis of a molecule of ATP. Note that the rear head, which was initially bound at site 1 (left panel), skips over the other head attached at site 2, and binds at site 3 to become the new leading head (right panel). The middle panels show the chemical and structural transitions that probably drive the motion: the binding of ATP^{57,60} (T) to the bound head is thought to produce a rotation in this head that swings the second head towards its next binding site³⁶. The second head then binds and releases its ADP (D). In order to be processive, it is essential that kinesin does not dissociate until after this second head is bound; this could be achieved if the tight binding associated with ADP release were to give the kick necessary to dislodge the first, now-trailing head. Perhaps this is the role for the tilting motion seen by cryoelectron microscopy upon ADP dissociation⁶⁸. The

high fidelity of the coordination would be maintained if the hydrolysis were contingent upon dissociation of the first head. If the two-heads-bound state is highly strained, then it is likely to be short-lived and can be considered as a transition state. In this case, the motor spends most of its time with just one head bound, and therefore the duty ratio is ~0.5. **b**, The motion of skeletal muscle myosin along the actin filament surface. In contrast to the microtubule, the two 'protofilaments' of the actin filament twist around the filament axis; as a result, myosin's consecutive binding sites are separated by $d = 36$ nm, and are on different protofilaments. Note that whereas kinesin's working stroke is large enough to span the 8 nm to the next binding site, a single myosin head cannot span the 36 nm to its next binding site: as a result the collaboration of at least ~10 myosin crossbridges in a macromolecular assembly is required to reach the next binding site. For convenience, only one head is shown—it is possible that the other head runs along a different actin filament. The exchange of ADP (D) for ATP (T) causes the head to unbind from the actin filament. The rebinding catalyses the release of phosphate (P) which initiates the working stroke.

biochemical constraints. If there were no steric limitations, then it would be expected that by decreasing the ATP concentration and therefore increasing the attached time, τ_{on} (see below), the number of attached crossbridges would increase (by equation (1)) and the force should increase. But the force actually decreases a little⁵⁴. Thus, in a muscle fibre only a small fraction of the myosin heads can be attached at the one time in force-generating states, even at very low ATP concentration, consistent with there being a steric constraint to binding.

Transitions that drive the working stroke

Identifying the chemical and conformational changes that drive the working stroke remains a central aim in the motor field. Solving the atomic structure of skeletal muscle myosin was a tremendous step forward. It revealed that the regulatory domain contains an 8-nm-long α -helix which is stabilized by two calmodulin-like light chains⁸, suggesting that this domain acted as a 'lever arm' that might amplify a much smaller conformational change in the nucleotide-binding pocket⁴¹. For example, a 30° rotation would provide the requisite 4 nm working distance. As reviewed in refs 55 and 56, the 'swinging lever-arm' hypothesis gains support from *in vitro* motility experiments showing that the speed is proportional to the length of the arm, and from electron microscopic and spectroscopic evidence for a swing. According to the standard biochemical model²⁰ shown in Fig. 3b, the rotation of the lever is driven by the release of phosphate from the actin–myosin–ADP state. The on time (τ_{on}) is determined primarily by the lifetime of this state: in a contracting muscle the lifetime is only about 1 ms, but this is long enough for the swing to be completed. After ADP dissociates, ATP binds and rapidly dissociates the actin–myosin complex. It is during the long-lived myosin–ATP and myosin–ADP–P_i states (which determine τ_{off}) that the motor is detached; but the actin filament continues to move owing to the action of other motors bringing the first motor to its next binding site, a multiple of 36 nm away, so that another cycle can begin. Thus the motor is carried through the full path distance even though it has contributed only about one tenth the distance with its own swing.

For kinesin, the postulated working stroke of 8 nm leads to a problem. How can the kinesin head, shown by crystallography to be only ~8 nm across, possibly span the 8 nm to the next binding site? One hypothesis is that kinesin's two heads undergo a coordinated motion so that movement within the attached head is amplified by the second, detached head^{35,36,57}—in this way the second head acts as the 'lever arm', rather like the regulatory domain of myosin. This would account for the slow speeds of single-headed kinesin fragments⁵⁸ in which the proposed lever is absent. A coordinated motion of the two heads also accounts for processivity if the release of the trailing head is contingent upon the binding of the leading head, analogous to the 'hand-over-hand' climbing of a rope^{17,59}. Such an alternating-head model is shown in Fig. 3a. The key idea is that a conformational change associated with ATP binding to the bound head^{57,60} leads to the motion of the second head so that it can bind to the next dimer and release its bound ADP. In this scheme, τ_{on} corresponds primarily to the microtubule–kinesin–ATP state and τ_{off} to the kinesin–ADP state: the coordination implies that for each head $\tau_{\text{on}} \cong \tau_{\text{off}}$, giving a duty ratio of 0.5.

The outstanding questions for both kinesin and myosin are how conformational changes associated with nucleotide hydrolysis generate force, and conversely, how mechanical force influences the rate constants of the chemical steps. For example, both myosin and kinesin slow down and eventually stall when the load is increased^{12,20,21,61}. If the duty ratio is to remain bounded by the mechanical constraints as discussed, then for equation (2) to remain valid at high external force, one or more of the following occurs: the ATPase rate decreases (Fenn effect²⁰); the effective working distance decreases owing to premature dissociation (slippage^{11,25}); or the stoichiometry of the coupling of the ATPase to the mechanical steps

decreases²¹. This is the next set of questions to be answered by single-molecule techniques.

Biological implications of the duty ratio

Let us now return to our original question, which was to see whether there are design principles that have emerged from the study of conventional kinesin and muscle myosin that can be used to understand the structures, functions and mechanisms of the more recently discovered motor proteins.

I have argued that the sizes of the working stroke and the distance between consecutive binding sites on the filament place a constraint on the biochemistry of motors—the fraction of time that a motor spends bound can be no larger than the ratio of the working distances and the path distances. Because the binding-site separation is so large, about 36 nm, the head of skeletal muscle myosin can spend no more than about 10–20% of its time bound to the actin filament, so the duty ratio is small. On the other hand, the relative proximity of the binding sites for kinesin, 8 nm, allows kinesin a high duty ratio; an obvious functional adaptation of a high duty ratio is that a single motor molecule suffices to transport organelles. Is there a functional reason why muscle myosin has a low duty ratio? A possible adaptation is for speed. Skeletal muscle myosin must be able to respond quickly if the load it is pulling on begins to yield: a low duty ratio ensures that there is a pool of primed heads that can come 'on line' after a small shortening as previously unreachable actin binding sites come into striking range. The muscle does not have to wait for all the heads to go around another cycle before it begins to shorten.

This is not to say that the biochemical rates cannot be tuned according to a motor's function. For example, the duty ratio of maximally shortening skeletal muscle is only about 0.01 (the stiffness is one fifth that of isometric muscle⁶², implying that the fraction of attached heads is only one tenth that of isometric muscle, considering the actin filament compliance mentioned above). Presumably the rebinding rate is so slow and the muscle contracting so fast that the heads are jumping over as many as ten potential binding sites. This is clearly an adaptation for higher efficiency because low loads require little work, and so a low duty ratio minimizes the number of ATP-consuming cycles. The differential tuning of biochemical rates can also explain the differences between axonemal dynein, which appears to have a low duty ratio⁶³, and cytoplasmic dynein, which ought to have a high duty ratio if it is to function as an organelle motor⁶⁴. If the duration of the detached state of axonemal dynein were ten times that of cytoplasmic dynein, then the duty ratio would be one tenth that of cytoplasmic dynein, explaining the approximately tenfold increase in speed of the axonemal over the cytoplasmic motor (Table 1). Likewise, it would seem fairly easy to make a non-processive kinesin, as might be the case for the kinesin-related protein Ncd⁶⁵, which functions in a large assembly—the meiotic spindle.

Finally it must be considered whether myosin could be made processive. This is particularly relevant to the myosin Vs, which are thought to act as vesicle transporters in yeast⁶⁶ and mice⁶⁷. To increase the duty ratio above 0.1–0.2, it is necessary to increase the working stroke or to decrease the binding-site separation. The only way to decrease the binding site separation would be for the motor to follow a different path but, as argued above, this would lead to spiralling of the organelle cargo around the actin filament, which is untenable as actin filaments are usually found close together. On the other hand, there are at least two ways of increasing the working stroke. One is to lengthen the neck. This is precisely what has happened with myosin V: each of its two heads has six light chains and a neck that is seen to be ~25 nm long by electron microscopy⁶⁷ (Fig. 4). This is almost long enough to reach the next actin binding site. So duty ratio considerations would say that the long neck is for processivity rather than for high speed—indeed, myosin V is a slow motor *in vitro* and *in vivo* (Table 1). Several other

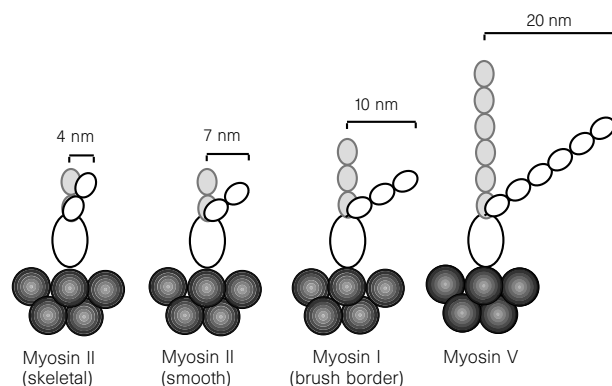


Figure 4 Ways in which myosin might increase its working stroke. By increasing the angle of the swing from $\sim 30^\circ$ to $\sim 60^\circ$, it is possible to increase the working distance from as little as ~ 4 nm for skeletal muscle myosin II (left), to ~ 7 nm for smooth muscle myosin II. By increasing the length of the lever arm by increasing the number of light chains in the regulatory domain to three or six, it is possible further to increase the working distance to ~ 10 nm for brush-border

myosin I, and to ~ 20 nm for myosin V. The larger the working distance, the larger the duty ratio and, as argued in the text, the fewer crossbridges are required to reach the next binding site. In the extreme case of myosin V, it is possible that just two heads in a dimer could reach to the next actin binding site 36 nm away, giving myosin V the capacity for processivity (as might be expected, given its biological function as a vesicle transporter).

members of the myosin superfamily also have 4–6 light chains, such as myosin VII, IX and XI, giving them potential for higher duty ratios. A second way to increase the working stroke is to increase the angle of the swing (Fig. 4): both smooth-muscle myosin⁶⁸ and brush-border myosin⁶⁹ Ib have an additional swing of about 30° associated with ADP release, bringing the total swing up to $\sim 60^\circ$. This could double the working distances of smooth muscle myosin to ~ 8 nm and myosin Ib to ~ 12 nm (as it has three light-chain domains). Such a mechanism explains the higher force generated by smooth-muscle myosin over skeletal-muscle myosin: the larger the working distance, the higher the duty ratio, and therefore, as postulated when the duty ratio concept was first formulated²⁴, the greater the fraction of the heads that are attached and generating force. An argument has recently been made against this mechanism⁴⁴; however, as mentioned earlier, these workers may have misinterpreted their single-molecule data.

Discussion

We have learnt a lot in the past several years about how motor proteins work. This is partially due to the new information that has come from the atomic structures of actin, myosin and kinesin, as well as the development of new single-molecule techniques that are beginning to revolutionize our understanding of molecular motors just as the patch clamp technique has revolutionized our understanding of ion channels⁷⁰. But perhaps as important has been the insight obtained by comparing the structures and functions of kinesin and myosin. In a sense, they have acted as controls for each other: our confidence that myosin is a low-duty-ratio motor is greatly increased by knowledge of how a high-duty-ratio motor like kinesin behaves in biochemical and motility assays; and whereas kinesin has given a clearer insight into the role of the filament and its discrete binding sites, myosin has given us a clearer look at the working stroke itself.

The unifying concept to emerge from this comparison of conventional kinesin and muscle myosin is that of the duty ratio—the fraction of the time that the motor spends in its attached, working state. The fundamental link between the structure and function of motor proteins is that the duty ratio can be no larger than the fraction of the distance that the working stroke moves the motor towards its next binding site. Understanding these constraints then gives insight into the structure and function of other motors, and determines whether the motor can move independently or whether motility is contingent on oligomerization into a macromolecular assembly. The low duty ratio of muscle myosin explains why it must

function in large macromolecular assemblies, while the high duty ratio of kinesin permits this motor to function on its own to transport organelles. The different duty ratios are not simply due to different biochemical rate constants governing attachment and detachment; instead, the duty ratios are constrained by the sizes of the conformational changes that the motors make, as well as the paths that the motors tread on the surfaces of their respective filamentous tracks.

The biochemistry can be tuned to decrease the duty ratio, and this may prove important to increase the speed of skeletal muscle over smooth muscle, or ciliary beating over organelle transport. On the other hand, the structure can be tuned to increase the duty ratio by lengthening the lever arm or increasing its swing; with a larger working stroke the distance to the next binding site can be more easily spanned. Application of single-molecule and protein engineering techniques to the diverse family of molecular motors promises to reveal a trove of design principles that might one day prove useful for building our own little machines. □

Joe Howard is in the Department of Physiology and Biophysics, University of Washington, Box 357290, Seattle, Washington 98195–7290, USA.

- Bloom, G. S. & Endow, S. A. Motor proteins 1: kinesins. *Protein Profile* **2**, 1105–1171 (1995).
- Barton, N. R. & Goldstein, L. S. B. Going mobile: microtubule motors and chromosome segregation. *Proc. Natl Acad. Sci. USA* **93**, 1735–1742 (1996).
- Mooseker, M. S. & Cheney, R. E. Unconventional myosins. *Annu. Rev. Cell Dev. Biol.* **11**, 633–675 (1995).
- Sellers, J. R. & Goodson, H. V. Motor proteins 2: myosin. *Protein Profile* **2**, 1323–1423 (1995).
- Gibbons, B. H., Asai, D. J., Tang, W.-J. Y., Hays, T. S. & Gibbons, I. R. Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchins. *Mol. Biol. Cell* **5**, 57–70 (1994).
- Gillespie, P. G. Molecular machinery of auditory and vestibular transduction. *Curr. Opin. Neurobiol.* **5**, 449–455 (1995).
- Bahler, M. Myosins on the move to signal transduction. *Curr. Opin. Cell Biol.* **8**, 18–22 (1996).
- Rayment, I. et al. Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–58 (1993).
- Kull, F. J., Sablin, E. P., Lau, R., Fletterick, R. J. & Vale, R. D. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature* **380**, 550–555 (1996).
- Svoboda, K., Schmidt, C. F., Schnapp, B. J. & Block, S. M. Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **256**, 721–727 (1993).
- Finer, J. T., Simmons, R. M. & Spudis, J. A. Single myosin molecule mechanics: piconewton forces and nanometer steps. *Nature* **368**, 113–119 (1994).
- Meyhöfer, E. & Howard, J. The force generated by a single kinesin molecule against an elastic load. *Proc. Natl Acad. Sci. USA* **92**, 574–578 (1995).
- Funatsu, T., Harada, Y., Tokunaga, M., Saito, K. & Yanagida, T. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* **374**, 555–559 (1995).
- Huxley, H. E. The mechanism of muscular contraction. *Sci. Am.* 18–27 (November 1965).
- Miller, R. H. & Lasek, R. J. Cross-bridges mediate anterograde and retrograde vesicle transport along microtubules in squid axoplasm. *J. Cell Biol.* **101**, 2181–2193 (1985).
- Sanger, J. M. & Sanger, J. W. Microtubule-based motility. *Cell Motil. Cytoskel.* (video suppl. 2, 1990).
- Howard, J., Hudspeth, A. J. & Vale, R. D. Movement of microtubules by single kinesin molecules. *Nature* **342**, 154–158 (1989).
- Harada, Y., Sakurada, K., Aoki, T., Thomas, D. D. & Yanagida, T. Mechanochemical coupling in actomyosin energy transduction studies by *in vitro* movement assay. *J. Mol. Biol.* **216**, 49–68 (1990).

19. Uyeda, T. Q. P., Kron, S. J. & Spudich, J. A. Myosin step size: estimation from slow sliding movement of actin over low densities of heavy meromyosin. *J. Mol. Biol.* **214**, 699–710 (1990).
20. Bagshaw, C. R. *Muscle Contraction* 2nd edn (Chapman and Hall, London, 1993).
21. Svoboda, K. & Block, S. M. Force and velocity measured for single kinesin molecules. *Cell* **77**, 773–784 (1994).
22. Molloy, J. E., Burns, J. E., Kendrick-Jones, J., Tregear, R. T. & White, D. C. S. Force and movement produced by a single myosin head. *Nature* **378**, 209–212 (1995).
23. Ishijima, A. *et al.* Multiple- and single-molecule analysis of the actomyosin motor by nanometer-piconewton manipulation with a microneedle: unitary steps and forces. *Biophys. J.* **70**, 383–400 (1996).
24. Dillon, P. F. & Murphy, R. A. High force development and cross-bridge attachment in smooth muscle from swine carotid arteries. *Circ. Res.* **50**, 799–804 (1982).
25. Leibler, S. & Huse, D. Porters versus crows: a unified stochastic model of motor proteins. *J. Cell Biol.* **121**, 1357–1368 (1993).
26. Toyoshima, Y. Y., Kron, S. J. & Spudich, J. A. The myosin step size. *Proc. Natl Acad. Sci. USA* **87**, 7130–7134 (1990).
27. Hackney, D. D. Highly processive microtubule-stimulated ATP hydrolysis by dimeric kinesin head domains. *Nature* **377**, 448–450 (1995).
28. Hua, W., Young, E. C., Fleming, M. L. & Gelles, J. Coupling of kinesin steps to ATP hydrolysis. *Nature* **388**, 390–393 (1997).
29. Schnitzer, M. J. & Block, S. Kinesin hydrolyses one ATP per 8-nm step. *Nature* **388**, 386–390 (1997).
30. Yanagida, T., Arata, T. & Oosawa, F. Sliding distance of actin filament induced by a myosin cross-bridge during one ATP hydrolysis cycle. *Nature* **316**, 366–369 (1985).
31. Higuchi, H. & Goldman, Y. E. Sliding distance between actin and myosin filaments per ATP molecule hydrolysed in skinned muscle fibres. *Nature* **352**, 352–354 (1991).
32. Wade, R. H. & Chrétien, D. Cryoelectron microscopy of microtubules. *J. Struct. Biol.* **110**, 1–27 (1993).
33. Ray, S., Meyhöfer, E., Milligan, R. A. & Howard, J. Kinesin follows the microtubule's protofilament axis. *J. Cell Biol.* **121**, 1083–1093 (1993).
34. Huang, T.-G., Suhan, J., Hackney, D. D. *Drosophila* kinesin motor domain extending to amino acid position 392 is dimeric when expressed in *Escherichia coli*. *J. Biol. Chem.* **269**, 16502–16507 (1994).
35. Hirose, K., Lockhart, A., Cross, R. A. & Amos, L. A. Three-dimensional cryoelectron microscopy of dimeric kinesin and *ncd* motor domains on microtubules. *Proc. Natl Acad. Sci. USA* **93**, 9539–9544 (1996).
36. Howard, J. The movement of kinesin along microtubules. *Annu. Rev. Physiol.* **58**, 703–729 (1996).
37. Sase, I., Miyata, H., Ishiwata, S. & Kinoshita, K. Jr Axial rotation of sliding actin filaments revealed by single-fluorophore imaging. *Proc. Natl Acad. Sci. USA* **94**, 5646–5650 (1997).
38. Suzuki, N., Miyata, H., Ishiwata, S. & Kinoshita, K. Jr Preparation of bead-tailed actin filaments: estimation of the torque produced by the sliding force in an *in vitro* motility assay. *Biophys. J.* **70**, 401–408 (1996).
39. Molloy, J. E., Burnes, J. E., Sparrow, J. C., Tregear, R. T., Kendrick-Jones, J. & White, D. C. S. Single-molecule mechanics of heavy meromyosin and S1 interacting with rabbit or *Drosophila* actins using optical tweezers. *Biophys. J.* **68**, 298s–305s (1995).
40. Eisenberg, E., Hill, T. L. & Chen, Y.-D. Crossbridge model of muscle contraction: quantitative analysis. *Biophys. J.* **29**, 195–227 (1980).
41. Rayment, I. *et al.* Structure of the actin–myosin complex and its implications for muscle contraction. *Science* **261**, 58–65 (1993).
42. Huxley, A. F. & Tideswell, S. Filament compliance and tension transients in muscle. *J. Musc. Res. Cell Motil.* **17**, 507–511 (1996).
43. Mehta, A. D., Fiber, J. T. & Spudich, J. A. Detection of single molecule interactions using correlated thermal diffusion. *Proc. Natl Acad. Sci. USA* **94**, 7927–7931 (1997).
44. Guilford, W. H. *et al.* Smooth muscle and skeletal muscle myosins produce similar unitary forces and displacements in the laser trap. *Biophys. J.* **72**, 1006–10021 (1997).
45. Block, S. M. One small step for myosin. *Nature* **378**, 132–133 (1995).
46. Huxley, A. F. & Simmons, R. M. Proposed mechanism for force generation in striated muscle. *Nature* **233**, 533–538 (1971).
47. Lombardi, V., Piazzesi, G. & Linari, M. Rapid regeneration of the actin–myosin power stroke in contracting muscle. *Nature* **355**, 638–641 (1992).
48. Huxley, H. E., Stewart, A., Sosa, H. & Irving, T. X-ray diffraction measurements of the extensibility of actin and myosin filaments in contracting muscle. *Biophys. J.* **67**, 2411–2421 (1994).
49. Wakabayashi, K. *et al.* X-ray diffraction evidence for the extensibility of actin and myosin filaments during muscle contraction. *Biophys. J.* **67**, 2422–2435 (1994).
50. Cooke, R., Crowder, M. S. & Thomas, D. D. Orientation of spin labels attached to cross-bridges in contracting muscle fibres. *Nature* **300**, 776–778 (1982).
51. Higuchi, H., Yanagida, T. & Goldman, T. E. Compliance of thin filaments in skinned fibers of rabbit skeletal muscle. *Biophys. J.* **69**, 1000–1010 (1995).
52. Kojima, H., Ishijima, A. & Yanagida, T. Direct measurement of stiffness of single actin filaments with and without tropomyosin by *in vitro* nanomanipulation. *Proc. Natl Acad. Sci. USA* **91**, 12962–12966 (1994).
53. Chen, Y.-D. & Brenner, B. On the generation of the actin–myosin power stroke in contracting muscle. *Proc. Natl Acad. Sci. USA* **90**, 5148–5152 (1993).
54. Ferenczi, M. A., Goldman, Y. E. & Simmons, R. M. The dependence of force and shortening velocity on substrate concentration in skinned muscle fibers from *Rana Temporaria*. *J. Physiol.* **350**, 519–543 (1984).
55. Block, S. Fifty ways to love your lever: myosin motors. *Cell* **87**, 151–157 (1996).
56. Holmes, K. C. The swinging lever-arm hypothesis of muscle contraction. *Curr. Biol.* **7**, R112–R118 (1997).
57. Hackney, D. D. Evidence for alternating head catalysis by kinesin during microtubule-stimulated ATP hydrolysis. *Proc. Natl Acad. Sci. USA* **91**, 6865–6869 (1994).
58. Stewart, R. J., Thaler, J. P. & Goldstein, L. S. B. Direction of microtubule movement is an intrinsic property of the motor domains of kinesin heavy chain and *Drosophila ncd* protein. *Proc. Natl Acad. Sci. USA* **90**, 5209–5213 (1993).
59. Schnapp, B. J., Crise, B., Sheetz, M. P., Reese, T. S. & Khan, S. Delayed start-up of kinesin-driven microtubule gliding following inhibition by adenosine 5'-[β - γ -imido]triphosphate. *Proc. Natl Acad. Sci. USA* **87**, 10053–10057 (1990).
60. Ma, Y.-Z. & Taylor, E. W. Interacting head mechanism of microtubule–kinesin ATPase. *J. Biol. Chem.* **272**, 724–730 (1997).
61. Hunt, A. J., Gittes, F. T. & Howard, J. The force exerted by a kinesin molecule against a viscous load. *Biophys. J.* **67**, 766–781 (1994).
62. Ford, L. E., Huxley, A. F. & Simmons, R. M. Tension transients during steady shortening of frog muscle fibers. *J. Physiol.* **361**, 131–150 (1985).
63. Hamasaki, T., Holwill, M. E., Barkalow, K. & Satir, P. Mechanochemical aspects of axonemal dynein activity studied by *in vitro* microtubule translocation. *Biophys. J.* **69**, 2569–2579 (1995).
64. Paschal, B. M., Shpetner, H. S. & Vallee, R. B. MAP 1C is a microtubule-activated ATPase which translocates microtubules *in vitro* and has dynein-like properties. *J. Cell Biol.* **105**, 1273–1282 (1987).
65. Crevel, I. M.-T. C., Lockhart, A. & Cross, R. A. Kinetic evidence for low chemical processivity in *ncd* and Eg5. *J. Mol. Biol.* (in the press).
66. Govindan, B., Bowser, R. & Novick, P. The role of Myo2, a yeast class V myosin, in vesicular transport. *J. Cell Biol.* **128**, 1055–1068 (1995).
67. Cheney, R. E. *et al.* Brain myosin-V is a two-headed unconventional myosin with motor activity. *Cell* **75**, 13–23 (1993).
68. Whittaker, M. *et al.* A 35-Å movement of smooth muscle myosin on ADP release. *Nature* **378**, 748–751 (1995).
69. Jontes, J. D., Wilson-Kubalek, E. M. & Milligan, R. A. A 32° tail swing in brush border myosin I on ADP release. *Nature* **378**, 751–753 (1995).
70. Hille, B. *Ionic Channels of Excitable Membranes* 2nd edn (Sinauer, Sunderland, MA, 1992).
71. Zot, H. G., Doberstein, S. K. & Pollard, T. D. Myosin-I moves actin filaments on a phospholipid substrate: implications for membrane targeting. *J. Cell Biol.* **116**, 367–376 (1992).
72. Ostap, E. M. & Pollard, T. D. Biochemical kinetic characterization of the Acanthamoeba myosin-I ATPase. *J. Cell Biol.* **132**, 1053–1060 (1996).
73. Toyoshima, Y. Y. *et al.* Myosin subfragment-1 is sufficient to move actin filaments *in vitro*. *Nature* **328**, 536–539 (1987).
74. Siemankowski, R. F., Wiseman, M. O. & White, H. D. ADP dissociation from actomyosin subfragment 1 is sufficiently slow to limit the unloaded shortening in vertebrate muscle. *Proc. Natl Acad. Sci. USA* **82**, 658–662 (1985).
75. Warshaw, D. M., Derosiers, J. M., Work, S. S. & Trybus, K. M. Smooth muscle myosin cross-bridge interactions modulate actin filament sliding velocity *in vitro*. *J. Cell Biol.* **111**, 453–463 (1990).
76. Rivolta, M. N., Urrutia, R. & Kachar, B. A soluble motor from the alga *Nitella* supports fast movement of actin filaments *in vitro*. *Biochim. Biophys. Acta* **1232**, 1–4 (1995).
77. Brennen, C. & Winet, H. Fluid mechanics of propulsion by cilia and flagella. *Annu. Rev. Fluid Mech.* **9**, 339–398 (1977).
78. Yokota, E. & Mabuchi, I. C. A dynein isolated from sea urchin sperm flagellar axonemes. Enzymatic properties and interaction with microtubules. *J. Cell Sci.* **107**, 353–361 (1994).
79. Brady, S. T., Pfister, K. K. & Bloom, G. S. A monoclonal antibody against kinesin inhibits both anterograde and retrograde fast axonal transport in squid axoplasm. *Proc. Natl Acad. Sci. USA* **87**, 1061–1065 (1990).
80. Shpetner, H. S., Paschal, B. M. & Vallee, R. B. Characterization of the microtubule-activated ATPase of brain cytoplasmic dynein (MAP 1C). *J. Cell Biol.* **107**, 1001–1009 (1988).
81. Kozminski, K. G., Beech, P. L. & Rosenbaum, J. L. The *Chlamydomonas* kinesin-like protein FLA10 is involved in motility associated with the flagellar membrane. *J. Cell Biol.* **131**, 1517–1527 (1995).
82. Cole, D. G. *et al.* Novel heterotrimeric kinesin-related protein purified from sea urchin eggs. *Nature* **366**, 268–270 (1993).
83. Yeh, E., Skibbens, R. V., Cheng, J. W., Salmon, E. D. & Bloom, K. Spindle dynamics and cell cycle regulation of dynein in the budding yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* **130**, 687–700 (1995).
84. Saunders, W. S. & Hoyt, M. A. Kinesin-related proteins required for structural integrity of the mitotic spindle. *Cell* **70**, 451–458 (1992).
85. Lockhart, A. & Cross, R. A. Kinetics and motility of the Eg5 microtubule motor. *Biochemistry* **35**, 2365–2373 (1996).
86. Chandra, R., Salmon, E. D., Erickson, H. P., Lockhart, A. & Endow, S. A. *J. Biol. Chem.* **268**, 9005–9013 (1993).
87. Lockhart, A., Cross, R. A. & McKillop, D. F. A. ADP release is the rate-limiting step of the MT-activated ATPase of non-claret disjunctional and kinesin. *FEBS Lett.* **368**, 531–535 (1995).
88. Hirose, K., Lockhart, A., Cross, R. A. & Amos, L. A. Nucleotide-dependent angular change in kinesin motor domain bound to tubulin. *Nature* **376**, 277–279 (1995).

Acknowledgements. This review was written while the author was a Guggenheim fellow on sabbatical leave in the Abteilung Molekulare Zellforschung, Max-Planck-Institut für Medizinische Forschung, Heidelberg. Original research was supported by the NIH and the Human Frontier Science Program. I thank many colleagues in Seattle and Heidelberg as well as M. Geeves, H. Gutfreund and M. Irving for critical comments on earlier drafts of the manuscript.

Constraints on flux rates and mantle dynamics beneath island arcs from Tonga–Kermadec lava geochemistry

Simon Turner & Chris Hawkesworth

Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK

Subduction processes are central to plate tectonics and to crust–mantle recycling and differentiation. Here we present a study of lavas from the Tonga–Kermadec island arc which places important constraints on the processes and rates involved. The mantle wedge overlying the subducting oceanic plate is dynamically coupled to the descending plate, but may convect more slowly than expected. Fluid and sediment fluxes from the ocean plate enrich the wedge but differ in their location, mechanisms and rates. After partial melting, magma extraction occurs rapidly via channelled flow through the wedge.

A fundamental tenet of modern plate-tectonic theory is the subduction of oceanic plates into the mantle. Subduction induces convective overturn within the mantle wedge^{1,2} and may also localize the sheeted downwellings of upper-mantle convection³. The surface expressions of subduction are the curved chains of active volcanoes that form island arcs, and the magmatic flux in these volcanoes is an important component of new crustal additions. Conversely, the subduction of oceanic crust and sediments forms the principal mechanism for the recycling of crustal materials into the mantle, contributing to mantle heterogeneity⁴. Thus island arcs play a pivotal role in models for the geochemical evolution of the continental crust and upper mantle, and an understanding of the physical processes and flux rates involved is one of the main goals of the Earth sciences.

Constraints on the dynamics of subduction come from both numerical models and geochemical studies. Recent numerical models^{1,2,5} have suggested that convection is induced in the mantle wedge by viscous drag along the upper surface of the subducting oceanic plate and that this controls the thermal structure of the system. The rate of convection is presumably linked to the rate of subduction but the extent of this coupling is poorly constrained, so most models at present assume complete coupling. As we show here, geochemical tracers of specific sediments can be used to constrain the transport time of the subducted sediment signal and thereby provide information on convective overturn in the mantle wedge. It is generally accepted that partial melting and volcanism result from lowering of the peridotite solidus by addition of aqueous fluids from the dehydrating oceanic crust^{1,2,5}. However, the rate of fluid flux, its location and whether it occurs vertically⁵ or migrates laterally across the mantle wedge^{1,2} remain to be determined. Studies of U-series isotope disequilibria have the potential to elucidate the rates of mass transfer. This is because under oxidizing conditions U is mobile in aqueous fluids whereas Th is not^{6,7} and the timescales of resultant disequilibria are similar to those of fluid-transfer and melt-generation processes (see, for example, refs 8–13).

Backarc spreading and basalt extraction efficiently depletes the incompatible-element budget of the mantle wedge beneath many island arcs¹⁴. Consequently, the signatures of contributions from the downgoing slab, particularly ²³⁸U-excesses, are best developed in the most depleted arc lavas^{15–18}. The Tonga–Kermadec arc is arguably the best place worldwide in which to evaluate mantle wedge

dynamics and flux rates because the plate tectonic setting is well documented^{19,20} and the very high rates of backarc spreading²⁰ have made this an exceptionally depleted, end-member arc system²¹. Consistent with this, preliminary U-series isotope data have indicated very large disequilibria in some of the lavas from this arc^{16,17}. Here we present selected results from a detailed elemental and isotopic study of Tonga–Kermadec and explore the new constraints these place on the mantle wedge dynamics and flux rates beneath the arc.

Geological background

The intra-oceanic Tonga–Kermadec island arc (Fig. 1) has been formed in response to westward-directed subduction of Jurassic–Cretaceous aged Pacific oceanic lithosphere beneath the Indo-Australian plate since the Oligocene epoch. Splitting of the arc around 3–6 Myr ago initiated active backarc spreading and formation of the Lau basin and Havre trough behind the Tonga–Kermadec ridge^{19,20,22–24}. The Tonga and Kermadec portions of the arc are divided at the point of subduction of the Louisville ridge¹⁹ and the present-day arc has a crustal thickness of 12–18 km and lies on volcanic basement²⁵. Convergence and backarc spreading rates both increase northwards along the arc (Fig. 1 and ref. 20) and the dip of the Benioff zone is ~28–30° to a depth of ~100 km (the average depth beneath the arc volcanoes) after which it steepens to 55–60° beneath the Kermadecs and 43–45° beneath Tonga²⁶. The volcanic products consist of tholeiitic basalts to basaltic andesites with lesser dacites and rhyolites and have phenocryst equilibration temperatures of 1,230–1,120 °C (ref. 25).

Analytical results and implications

We selected 58 samples spanning the length of the arc; we also selected 19 samples of the sediments cored at Deep Sea Drilling Program (DSDP) Site 204 and lavas from the backarc island of Niuafu'ou which lies in the Lau basin some 200 km behind the northern Tongan arc (see Fig. 1). The full data set is presented elsewhere²⁷; selected results, pertinent to the arguments below, are given in Table 1.

The geochemical and isotopic data reveal that four components are involved in the petrogenesis of the Tonga–Kermadec lavas. On a ²⁰⁸Pb/²⁰⁴Pb–²⁰⁶Pb/²⁰⁴Pb diagram (Fig. 2a), most of the Tongan lavas overlap the field for the backarc basalts from the Lau basin²⁴ which are taken as representative of the mantle wedge composition. The

Kermadec lavas extend from this array towards elevated $^{208}\text{Pb}/^{204}\text{Pb}$. In marked contrast, the lavas from the northernmost Tongan islands of Tafahi and Niuatoputapu are strongly displaced to high $^{206}\text{Pb}/^{204}\text{Pb}$. The extremely depleted nature of the lavas is exemplified by the concentrations of high-field-strength elements (for example, Ta and Nb) and rare-earth elements (for example, Nd), which are 10 times lower than in typical mid-ocean-ridge basalts (MORBs) and are in fact more comparable with abundances in the MORB source. The large-ion lithophile elements (for example, Ba) are enriched relative to the high-field-strength elements and rare-

earth elements (for example, $\text{Ba}/\text{Ta} = 1,465\text{--}10,774$) although absolute concentrations remain very low. This indicates that the large-ion lithophile elements have been decoupled from the high-field-strength elements and rare-earth elements; this is widely interpreted in terms of a flux of large-ion lithophile elements from the subducting slab into a mantle wedge that was variously depleted in incompatible elements during multi-stage backarc basalt extraction^{21,27}. Therefore, one of the surprising results is that Ta and Nb abundances are not lowest in the most- depleted rocks from Tafahi and Niuatoputapu where $\text{Ba}/\text{Ta} = 1,465\text{--}10,774$.

Table 1 Selected data for Tonga–Kermadec lavas

Island	Mg (wt%)	Ba (p.p.m.)	Th (p.p.m.)	U (p.p.m.)	Ta (p.p.m.)	Nd (p.p.m.)	$^{87}\text{Sr}/^{86}\text{Sr}$	$^{143}\text{Nd}/^{144}\text{Nd}$	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$	$(^{230}\text{Th}/^{232}\text{Th})$
Tafahi	7.47	33	0.096	0.051	0.02	1.5	0.70388	0.512930	18.978	38.560	1.215
Niuatoputapu	3.31	120	0.399	0.174	0.06	4.3	0.70397	0.512874	19.020	38.720	1.195
Fonualei	1.54	206	0.487	0.472	0.05	7.6	0.70376	0.512992	18.533	38.141	1.652
Late	3.29	165	0.236	0.206	0.03	5.5	0.70369	0.512942	18.518	38.125	1.603
Metis Shoal	5.15	362	0.308	0.230	0.04	5.6	0.70348	0.512980	18.473	37.986	1.821
Kao	5.56	123	0.188	0.140	0.03	6.0	0.70328	0.513068	18.454	37.995	1.401
Tofua	6.16	95	0.132	0.119	0.01	3.0	0.70370	0.513020	18.527	38.098	1.645
Hunga Ha'apai	4.16	119	0.121	0.116	0.03	3.0	0.70371	0.513040	18.516	38.148	1.799
Ata	6.72	102	0.266	0.136	0.02	4.6	0.70344	0.513087	18.686	38.240	1.335
Raoul	4.82	87	0.235	0.164	0.02	5.7	0.70341	0.513030	18.686	38.371	1.273
Macauley	5.45	119	0.438	0.231	0.02	8.1	0.70343	0.513024	18.675	38.306	1.144
L'Esperance	4.87	121	0.460	0.155	0.03	5.5	0.70394	0.512951	18.729	38.427	0.874
Rumble III	6.47	230	0.416	0.176	0.04	8.8	0.70409	0.513032	18.721	38.538	1.016
Niuao'ou	7.16	44	0.309	0.095	0.29	12.3	0.70432	0.512834	18.266	38.202	1.026

Methods of analysis: MgO, X-ray fluorescence; Ba, Ta and Nd, inductively coupled plasma mass spectrometry; Th and U, and isotopes of Nd, Pb and Th isotopes, thermal ionization mass spectrometry (see refs 9 and 27 for details).

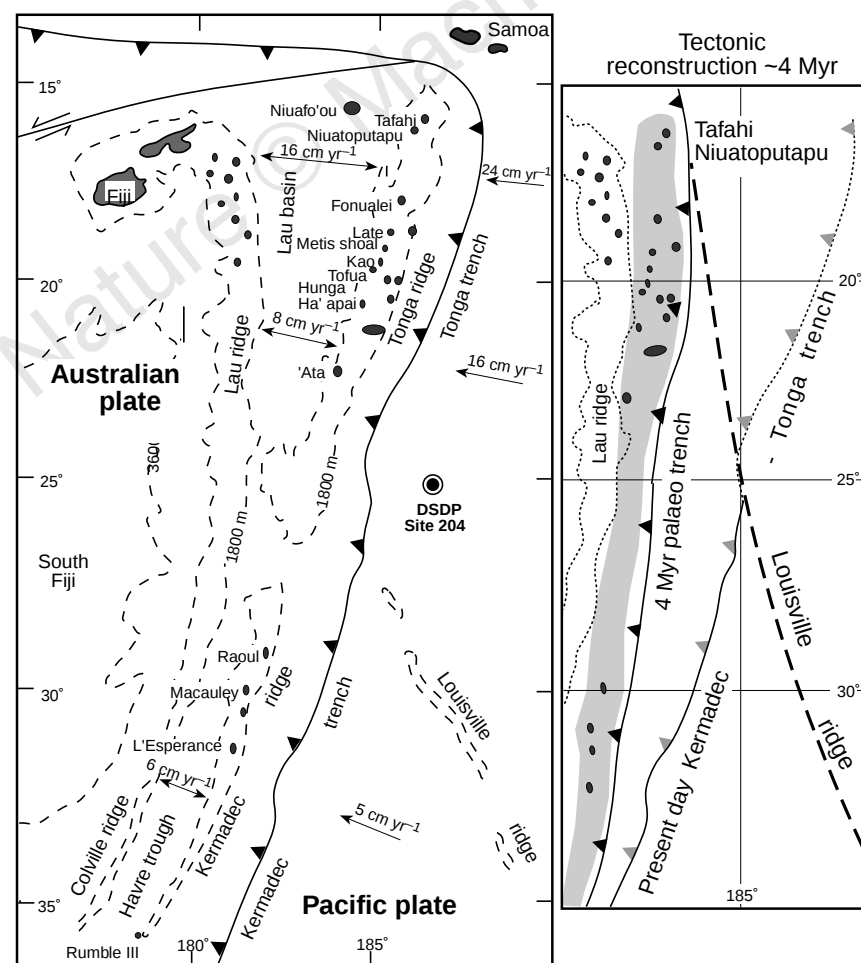


Figure 1 Map of the Tonga–Kermadec island arc system. This figure shows the bathymetry, locations of islands from which samples analysed in this study were taken, and the broad tectonic architecture with estimated plate motions indicated. Convergence and backarc spreading rates both increase northwards along the arc from 50 to 240 mm yr^{-1} and 60 to 160 mm yr^{-1} respectively²⁰. DSDP Site 204 from which the subducting sediments have been analysed in detail²⁷ is also shown (sites 595/596 are ~1,000 km further east). The angle between the Louisville ridge and the arc (which is broadly orthogonal to the motion of the Pacific plate) causes the locus of intersection to migrate southwards with time; the right-hand panel is a tectonic reconstruction¹⁹ showing that the Louisville ridge and its surrounding apron of volcanoclastic sediments would have been being subducted beneath the northern Tongan islands of Tafahi and Niuatoputapu ~4 Myr ago.

In Fig. 2b most of the lavas form an array in which decreases in $^{143}\text{Nd}/^{144}\text{Nd}$ are accompanied by constant, or very slightly decreasing, Ta/Nd . Mass-balance calculations^{9,10,27–29} have shown that addition of even very small amounts of subducted sediment to the source of the arc lavas will dominate the final Th, Ta–Nb, Zr and rare-earth-element budget of the lavas and consequently dictate their Nd- and Th-isotope signatures^{9,10,13,27,28} (isotopes of Sr and Pb reflect a more complex balance between the relative compositions and contributions from the fluid and sediment^{9,27,30,31}). Thus, the low Ta/Nd and $^{143}\text{Nd}/^{144}\text{Nd}$ coupled with elevated $^{208}\text{Pb}/^{204}\text{Pb}$ are consistent with a sediment contribution to the lavas (see below). But the Tafahi and Niuatoputapu lavas, which were displaced to high $^{206}\text{Pb}/^{204}\text{Pb}$ in Fig. 2a, show a trend of increasing Ta/Nd with decreasing $^{143}\text{Nd}/^{144}\text{Nd}$ (Fig. 2b). Such a trend is extremely unusual in arc rocks and these lavas must be influenced by a component which is not seen elsewhere in the arc. Finally, Fig. 2c shows that the lavas form a steep negative array between Ba/Th and $^{206}\text{Pb}/^{204}\text{Pb}$ which does not fall on any simple mixing line between the Lau basin mantle and subducted sediment; this emphasizes the need for a fluid component with high Ba/Th (refs 9, 13, 28). The implied Pb isotopic signature of this fluid is significantly lower than that of the subducting sediment (Fig. 2c) and so the fluids are inferred to be derived by dehydration of the subducting altered oceanic crust rather than the subducted sediments^{9,13,30}.

Any model for the Tonga–Kermadec data therefore requires four components—the depleted mantle wedge, separate fluid and sedi-

ment contributions^{9,10,13,28}, and a fourth component characterized by high Ta/Nd and $^{206}\text{Pb}/^{204}\text{Pb}$ and low $^{143}\text{Nd}/^{144}\text{Nd}$ seen only in lavas from Tafahi and Niuatoputapu. Our U–Th isotope results are shown in Fig. 3 and include the most extreme ratios yet reported, with up to 50% ^{238}U excesses. The lavas from the backarc island of Niuafu’ou show no evidence for a contribution from the fluid or sediment components identified in the arc lavas (Fig. 2) and lie to the left of the equiline in Fig. 3 with 10–30% ^{230}Th excesses similar to those observed in MORB (see, for example, ref. 32).

Sediment flux

The sediments being subducted beneath the Tonga–Kermadec arc on the downgrowing Pacific plate have been intersected at DSDP Site 204 (Fig. 1) and sites 595/596 located ~1,000 km east of the trench. The total thickness is ~150 m and this is dominantly composed of siliceous ooze and clay throughout sites 595/596 and in the upper 103 m of Site 204. The average composition of this pelagic material is broadly similar to that of average upper crust^{29,27,33} and forms an appropriate end-member for the sediment-dominated trend observed in most of the lavas in Fig. 2a, b (that is, towards high $^{208}\text{Pb}/^{204}\text{Pb}$ and low Ta/Nd , $^{143}\text{Nd}/^{144}\text{Nd}$). The mixing line between the Lau basin mantle and subducted pelagic sediment in Fig. 2b suggests that the amount added is <1%, implying that most of the subducted sediment is recycled into the upper mantle.

In contrast, the basal 44 m from DSDP Site 204 is dominated by

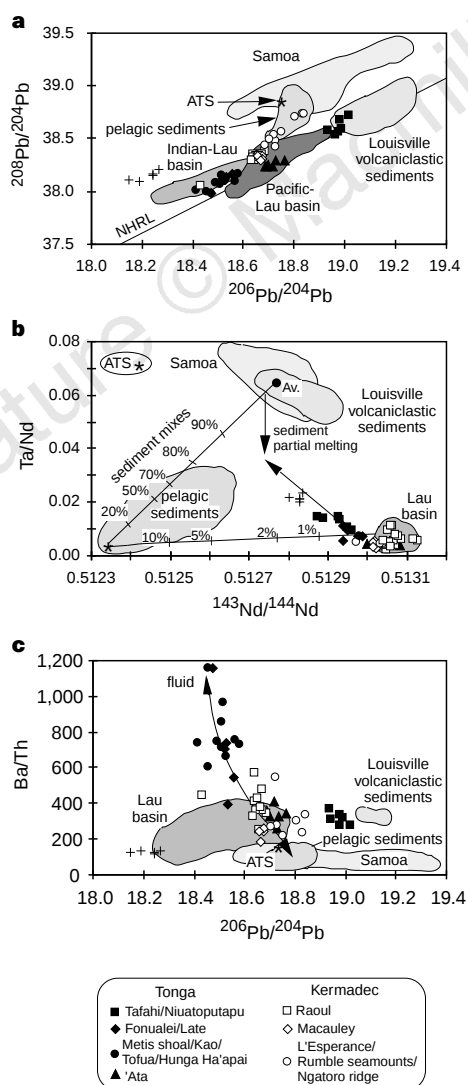


Figure 2 Geochemical and isotopic properties of the Tonga–Kermadec lavas. **a**, $^{208}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram showing the distinction between the average pelagic sediment (ATS)³³ being subducted beneath the arc and the Louisville volcanoclastic-dominated sediments from the base of DSDP 204 (ref. 27). The Tonga–Kermadec lavas lie on a mixing array between mantle wedge (represented by the basalt array from the Lau basin²⁴) and the pelagic sediments, excepting those from Tafahi and Niuatoputapu which are strongly displaced towards the Louisville volcanoclastics but not Samoa (data from refs 50 and 51). NHRL is the northern hemisphere reference line. **b**, Ta/Nd versus $^{143}\text{Nd}/^{144}\text{Nd}$ diagram, showing that most of the Tonga–Kermadec lavas lie on a mixing line (numbers along the curves indicate the amounts of sediment) between the Lau basin mantle and the average subducted pelagic sediment. In contrast, the Tafahi and Niuatoputapu lavas lie on a different vector displaced to high Ta/Nd and the sediment contribution to these lavas is implied to be a ~10:90 mixture of pelagic and Louisville volcanoclastic sediments. In practice, because of the high Pb content of the pelagic sediments, such a mixture would have a $^{206}\text{Pb}/^{204}\text{Pb}$ ratio of ~18.91 which is too low to explain the Tafahi–Niuatoputapu Pb isotope data (compare **a**). This constraint is relaxed if the sediment mixture was closer to 99% Louisville volcanoclastics. However, such a mixture will have a Ta/Nd ratio which is too high to lie at the end of the Tafahi–Niuatoputapu trend. Thus, it is suggested that the sediment signature was transferred as a partial melt (see also refs 9, 10, 27), such that its Ta/Nd ratio was reduced from ~0.6 to ~0.35 owing to the presence of residual phases like rutile or ilmenite⁴⁷ which preferentially held back Ta and Nb relative to the large-ion lithophile elements and rare-earth elements (see, for example, ref. 52). **c**, Ba/Th versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram showing a broad negative array that requires a fluid component as well as the mantle wedge and the subducted sediments. We note that the Tafahi and Niuatoputapu rocks are displaced from this trend towards the Louisville volcanoclastics. The crosses are data from the backarc island of Niuafu’ou.

Cretaceous volcanics derived from the nearby Louisville ridge³⁴ (see Fig. 1). Consistent with analyses of basalts from the Louisville seamount chain itself³⁵, these have an ocean-island basalt signature characterized by high Ta/Nd and high $^{206}\text{Pb}/^{204}\text{Pb}$ (ref. 27) providing a unique geochemical tracer which unveils the identity of the fourth component. Mixing arrays form straight lines on Pb–Pb isotope diagrams such as Fig. 2a which shows how a contribution bearing the Louisville signature accounts for the high $^{206}\text{Pb}/^{204}\text{Pb}$ signal in the Tafahi–Niuatoputapu lavas. A large volcanoclastic sediment apron also surrounds Samoa just to the north of the Tonga arc (see Fig. 1) and may extend as far south as Tafahi and Niuatoputapu³⁶, and so it might be argued that the high Ta/Nd and $^{206}\text{Pb}/^{204}\text{Pb}$ in those lavas reflects a contribution from subducted Samoan volcanoclastics. But because the Samoan plume has much higher $^{208}\text{Pb}/^{204}\text{Pb}$ than the Louisville volcanoclastics, mixing lines drawn between the Lau basin and Samoa will not pass through the Tafahi–Niuatoputapu data (Fig. 2a). Similarly in Fig. 2c, the Tafahi–Niuatoputapu rocks are displaced towards the Louisville volcanoclastics. Thus, it seems clear that the high Ta/Nd and high $^{206}\text{Pb}/^{204}\text{Pb}$ in the Tafahi–Niuatoputapu rocks are due to the addition of a Louisville component.

It is significant that, in detail, the data in Fig. 2b require that the contribution is a mixture of the pelagic sediments and the Louisville volcanoclastics. Owing to the age and velocity of the subducting oceanic plate, the Louisville seamounts themselves will be too cold to undergo partial melting³⁷, and so we conclude that it is the sediments themselves (rather than melts or fluids from the seamounts) that are providing the high $^{206}\text{Pb}/^{204}\text{Pb}$ and Ta/Nd signal. Furthermore, the data require the Ta/Nd ratio to be fractionated in the sediment component (Fig. 2b) and so it appears that the sediment component is added as partial melts that vein the mantle wedge peridotite²⁷.

The important point is that the Louisville volcanoclastics are spatially restricted to the proximity of the Louisville ridge; they are not observed at DSDP sites 595/596, for example. This is critical to

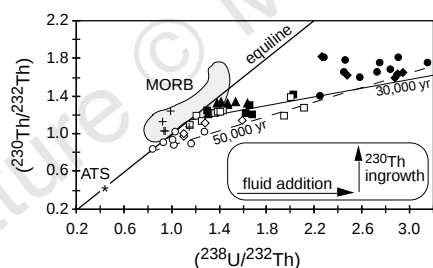


Figure 3 $(^{230}\text{Th}/^{232}\text{Th})$ versus $(^{238}\text{U}/^{232}\text{Th})$ (parentheses indicate activity ratios) 'equiline' diagram showing that the Tonga–Kermadec lavas with the largest U-excesses lie on or above a ~50,000-yr reference line (dashed). The simplest interpretation is that this reflects the typical time elapsed between fluid release from the slab and eruption. We attribute age significance to a reference line along the base of the data in preference to an isochron regression though the data because some of the transit times may be longer for some lavas which would result in increases in ^{230}Th and individual samples moving upwards from the reference line. We also note that if the sediment contribution to the Kermadec lavas has lowered their $(^{230}\text{Th}/^{232}\text{Th})$ intersection with the equiline, then 50,000 yr is probably a maximum estimate of the transfer time and a reference line beneath the Tonga data alone has an age of ~30,000 yr. (The inset illustrates our interpretation of the data.) We note that there is a broad negative correlation between $^{208}\text{Pb}/^{206}\text{Pb}$ (radiogenic lead) and $(^{230}\text{Th}/^{232}\text{Th})$ (ref. 27). Previously, such correlations have been used to argue that the significant, global U/Th variation among arcs is long-lived (that is, several 100 Myr; ref. 17). Within an individual and very depleted arc such as Tonga–Kermadec, such an array is interpreted to represent recent mixing of components (that is, sediment and mantle wedge + fluid) which have themselves had long-lived differences in their U/Th ratios, rather than to imply that the observed variations in U/Th along the arc had existed there for several 100 Myr; ref. 27. Symbols as in Fig. 2, field for MORB from ref. 32.

any evaluation of the time taken for such components to traverse the mantle wedge because the Louisville ridge at present intersects the arc 1,100 km south of Tafahi and Niuatoputapu. The angle between the Louisville ridge and the arc causes the locus of intersection to migrate southwards with time. The present-day convergence rate is 24 cm yr^{-1} but sea-floor spreading rates in the Lau basin have increased in the recent past³⁸. If the convergence rate was closer to 8 cm yr^{-1} at the start of backarc spreading 4 Myr ago, then an average convergence rate of $\sim 15\text{ cm yr}^{-1}$ may be more reasonable. Plate reconstructions on the basis of a convergence rate of 15 cm yr^{-1} indicate that the Louisville ridge and its apron of volcanoclastic sediments would have been subducted beneath Tafahi and Niuatoputapu ~4 Myr ago¹⁹ (Fig. 1); this probably represents a maximum age. A striking observation is that the Louisville signature is not observed on the islands south of Niuatoputapu (Fig. 2a, b) and the high $^{143}\text{Nd}/^{144}\text{Nd}$ ratios (0.51306–0.51308; ref. 22) of 4-Myr-old lavas from the northern Lau ridge, which formed part of the arc before backarc spreading, suggest that the Louisville signature was not present beneath Tafahi and Niuatoputapu 4 Myr ago (unfortunately no Pb isotope or reliable Ta data are available for the Lau ridge lavas). Taken together, and allowing for possible errors in the plate reconstructions, these observations imply that 2–4 Myr elapses between the time of subduction of the sediments and their signature being observed in the arc lavas.

Fluid flux

Experimental data have shown that U (as well as Ba, K, Sr and Pb) is highly mobile in oxidizing aqueous fluids whereas Th behaves as an immobile high-field-strength element^{6,7}. Redox conditions will be strongly oxidizing in the altered oceanic crust but reducing in the mantle wedge. Accordingly, the U excesses are taken to reflect the addition of fluids released by dehydration of the subducting oceanic crust^{9,10,13,16–18,27,28}. As interaction with the wedge proceeds, the fluids can only become more reducing under which circumstances U will become less mobile and the fractionation of U/Th will diminish. So, the high U/Th results from the initial fluid addition to the wedge and it is those samples with the least sediment contribution (inferred from Nd isotopes²⁸) which show the greatest U–Th isotope disequilibria. Such disequilibria are reduced over time by ^{230}Th ingrowth and the data can thus be used to constrain the time elapsed between fluid release from the slab and eruption. The Tonga–Kermadec data indicate that this was of the order of 30,000–50,000 yr (see Fig. 3). Significantly, U–Th isotope data record similar timescales in the Lesser Antilles (~40,000 yr; ref. 9) and in the Marianas (30,000 yr; ref. 10) which provides encouragement for the view that these data reflect some general aspect of the transfer times beneath island arcs.

As shown in Fig. 2c, Ba/Th is fractionated by the fluid and as Ra is likely to behave in an analogous manner to Ba, Ra/Th must also be fractionated by the fluid. However, ^{226}Ra will return to secular equilibrium with ^{230}Th within 7,500 yr of Ra/Th fractionation and well within the inferred time period (30,000–50,000 yr) between fluid release and eruption. Nevertheless, large ^{226}Ra excesses have been reported in several of the same samples analysed for U–Th from Tonga–Kermadec ($(^{226}\text{Ra}/^{230}\text{Th}) = 1.5\text{--}3.0$; ref. 16), the Lesser Antilles ($(^{226}\text{Ra}/^{230}\text{Th}) = 1.2\text{--}2.2$; ref. 11, 16) and the Marianas ($(^{226}\text{Ra}/^{230}\text{Th}) = 2.6\text{--}3.1$; ref. 16). Here the parentheses indicate activity ratios. This would not be expected given the ages inferred for the U–Th data, suggesting that the ^{238}U – ^{230}Th and ^{226}Ra – ^{230}Th disequilibria are decoupled (see also ref. 9). In other words, any Ra/Th fractionation produced during fluid release returns to secular equilibrium and the observed $^{226}\text{Ra}/^{230}\text{Th}$ disequilibria is developed subsequently.

Partial melting and magma extraction

The decoupling between the ^{238}U – ^{230}Th and ^{226}Ra – ^{230}Th disequilibria can be integrated into a model for partial melting and magma

extraction. Ba (and by analogy Ra) partitions more strongly into plagioclase than Th (ref. 39) and so it might be argued that the ^{226}Ra excesses result from plagioclase accumulation. However, the Tonga–Kermadec major-element data indicate that plagioclase is a fractionating phase and therefore the ^{226}Ra excesses are unlikely to have developed in crustal magma chambers. By analogy with MORB, we suggest that the ^{226}Ra excesses are produced during partial melting and/or magma ascent^{8,12,40}. In current melting models for mid-ocean ridges, Ra is not sensitive to the depth of melting and we infer that the ^{226}Ra excesses in arc lavas indicate melt extraction and ascent on timescales no more than 1–2 half lives of ^{226}Ra (that is, 1,600–3,200 yr) which requires channelled flow through the mantle wedge and ascent rates of $\sim 35\text{--}70\text{ mm yr}^{-1}$. Similar magma ascent rates have been proposed to occur beneath mid-ocean ridges⁸ and there is as yet little reason to suppose that they should be any slower beneath island arcs. In fact, recent numerical modelling suggests that the mantle wedge is likely to be under tension directly beneath the arc volcanoes⁴¹, and high volatile contents would result in low magma viscosities, both of which would promote rapid magma ascent.

Implications for physical models

As illustrated in Fig. 4, our data provide important constraints for physical models of the Tonga–Kermadec island arc and may be applicable to arcs in general.

(1) Recycling of the sediment component is slow; it takes 2–4 Myr after subduction before the Louisville volcanics signature is observed in the arc lavas. Yet the rate of subduction beneath Tonga is $150\text{--}240\text{ mm yr}^{-1}$ and so sediments would arrive at the source region beneath the arc volcanoes in $<1\text{ Myr}$ if they were transported on the subducting plate. In 2–4 Myr they would have passed several 100 km beyond the arc volcanoes, and yet there is no evidence for a sediment signature in lavas from the backarc island of Niufo’ou as would be expected if the sediment component was being brought into the arc melt generation zone via convection from the backarc. It might be suggested that the sediments travel on the descending plate to a point beneath the arc volcanoes where they undergo partial melting and that it subsequently takes 2–4 Myr for these melts to migrate upwards to the site of magma generation. However, partial melts of sediments are highly viscous and would freeze in the mantle

wedge⁴² after which they would be rapidly swept away by the convection in the wedge. The rate of convection in the mantle wedge is generally poorly constrained, but the contrast in velocities of the overlying and downgoing plates clearly demands that there is some decoupling between these plate motions and the mantle wedge.

In our preferred model, partial melts of the sediment are added to the overlying mantle wedge at a relatively shallow level, and the wedge is only partially coupled to the descending slab. As a result, convection and down-dip movement are much slower than the rate of slab descent and this enriched mantle only reaches the melt-generation zone after 2–4 Myr. Depending on the exact point of addition, this constrains the down-dip component of induced convection in the mantle wedge to be $20\text{--}40\text{ mm yr}^{-1}$. We note that this is a maximum range assuming that the sediment signal is imparted to the mantle wedge at shallow levels during subduction (see below), but it is significant that this is similar to the half-spreading rates in the Lau basin. In other words, the mantle convection rate seems to be more closely linked to the rate of motion of the overlying plate than to that of the downgoing plate against which there must therefore be a major component of shear-slip.

(2) In contrast, recycling of the fluid component is relatively fast; the U–Th data indicate only 30,000–50,000 yr has elapsed since fluid release from the slab, and clearly this must have occurred after addition of the sediment (see also ref. 10). Recent experimental data⁴³ suggest that subducted oceanic crust undergoes pressure-dependent amphibolite to eclogite dehydration reactions at around 70–80 km depth. In some models the resultant fluid flux occurs vertically into the wedge and the amphibole peridotite formed then migrates downwards with convection in the wedge until it crosses its solidus ($\sim 1,000^\circ\text{C}$; refs 44, 45) and undergoes partial melting⁵. Alternatively, it has been suggested that the fluid migrates horizontally across the mantle wedge by a combination of vertical rise as a fluid and horizontal translation whilst being dragged down in the wedge in the form of amphibole until it reaches the amphibole peridotite solidus^{1,2}. The latter model is consistent with the lack of evidence from the rare-earth elements and U–Th disequilibria that melting occurred in the presence of residual garnet. If the melt generation zone lies at 70–80 km depth, then, assuming that magma

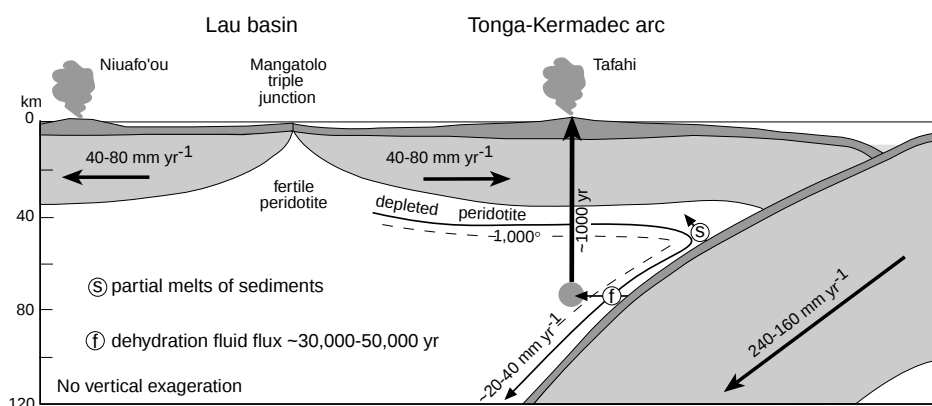


Figure 4 Scaled east-west cross section across the Tonga arc through Tafahi illustrating the inferred mantle wedge dynamics and flux rates beneath the Tonga–Kermadec island arc. The mantle wedge convects at a rate similar to that of the overlying plate and is largely decoupled from the downgoing plate. This produces significant shear-heating along its upper surface at shallow levels which, combined with advection of hot wedge material from the Lau basin, results in partial melting of the subducted sediments which enrich the depleted wedge peridotite at point S. This material is carried down with convection in the wedge for 2–4 Myr until point f where amphibolite in the subducted oceanic crust dehydrates, releasing aqueous fluids. These fluids traverse the wedge until the amphibole peridotite solidus ($1,000^\circ\text{C}$ isotherm^{44,45}) is reached and partial melting

of sediment enriched peridotite occurs. Rapid magma ascent to the surface then occurs via channelled flow. Neither the sediment or fluid components reach the source region for the backarc volcano of Niufo’ou. We note that the inferred transport time of $\sim 4\text{ Myr}$ for the sediment component would predict that the majority of any ^{10}Be signal from the subducted sediments would have decayed away in this time. Unfortunately, there are as yet no Be isotope data from Tonga–Kermadec to test this. Moreover, because of the condensed nature of the subducting sediment column in the Tonga–Kermadec arc, all of the ^{10}Be may be concentrated in the upper 10 cm and vulnerable to mechanical loss during subduction (see refs 13 and 27 for further discussion of the ^{10}Be signal in arc lavas).

ascent is more or less vertical, it must lie 20–30 km out from the slab such that the arc volcanoes lie ~100 km above the slab. But given the inferred rate of convection in the wedge, the angle of subduction and the 30,000–50,000 yr time constraints for fluid transport, the horizontal component of translation in the wedge is limited to ~1–3 km, suggesting that some other process, such as hydraulic fracture⁴⁶, is required to allow fluids to be transported more efficiently through the mantle wedge¹³.

(3) The data reported here seem to require a thermal structure within the mantle wedge that is considerably hotter than suggested by recent thermal models^{1,2} (this conclusion does depend, however, on the details of the models). Nevertheless, the evidence for partial melting of the subducted sediments (see also refs 9, 10) and the phenocryst equilibration temperatures of the lavas (1,120–1,230 °C) may support the existence of a hotter thermal structure. In numerical models the thermal structure of the wedge is least well defined near the slab–wedge interface², and at present there are few experimental studies of partial melting of hydrated sediments under pressure–temperature conditions appropriate to the sub-arc environment⁴⁷. However, sediment melting probably requires temperatures of 650–700 °C. The decoupling required by the relative plate motions and the transfer time of the Louisville component suggest that significant shear occurs against the downgoing plate and, although the effects of shear-heating are poorly constrained, this may increase temperatures by as much as 200 °C at shallow levels⁴⁸. Moreover, the presence of 1.4–2.0 Myr boninites dredged between Tafahi and the trench⁴⁹ provides evidence for high temperatures in the wedge close to the trench at that time. Irrespective of the model assumed, the limited time for fluid transport following dehydration of the oceanic crust at 70–80 km depth may even require that the position of the 1,000 °C isotherm and zone of melt generation lies within a few kilometres of the slab–wedge interface between 70 and 100 km depth.

(4) The sediment and fluid signatures are not observed in lavas from the backarc island of Niuafo'ou. This places a maximum limit on the lateral extent to which the subduction signature penetrates into the mantle wedge (<200 km). The Niuafo'ou rocks preserve ²³⁰Th excesses similar to MORBs (Fig. 3), consistent with partial melting in the presence of residual garnet and at deeper levels than beneath the arc volcanoes.

(5) Reported ²²⁶Ra disequilibria¹⁶ suggest that the U–Th and Ra–Th systematics are decoupled; the simplest explanation is that the observed ²²⁶Ra excesses developed during partial melting and magma ascent^{8,40} rather than during fluid release for the subducting slab. Thus the magmas, once formed, ascend rapidly (~1,600–3,200 yr) via channelled flow through the mantle wedge to be erupted along the overlying chain of arc volcanoes. □

Received 11 December 1996; accepted 29 July 1997.

- Davies, J. H. & Bickle, M. J. A physical model for the volume and composition of melt produced by hydrous fluxing above subduction zones. *Phil. Trans. R. Soc. Lond.* **335**, 355–364 (1991).
- Davies, J. H. & Stevenson, D. J. Physical model of source region of subduction zone volcanics. *J. Geophys. Res.* **97**, 2037–2070 (1992).
- Bercovici, D., Schubert, G. & Glatzmaier, G. A. Three-dimensional spherical models of convection in the Earth's mantle. *Science* **244**, 950–955 (1989).
- Hofmann, A. W. Mantle geochemistry: the message for oceanic volcanism. *Nature* **385**, 219–229 (1997).
- Tatsumi, Y. Migration of fluid phases and genesis of basalt magmas in subduction zones. *J. Geophys. Res.* **94**, 4697–4707 (1989).
- Brenan, J. M., Shaw, H. F., Ryerson, F. J. & Phinney, D. L. Mineral–aqueous fluid partitioning of trace elements at 900 °C and 2.0 GPa: constraints on the trace element chemistry of mantle and deep crustal fluids. *Geochim. Cosmochim. Acta* **59**, 3331–3335 (1995).
- Keppeler, H. Constraints on partitioning experiments on the composition of subduction-zone fluids. *Nature* **380**, 237–240 (1996).
- Richardson, C. & McKenzie, D. Radioactive disequilibria from 2D models of melt generation by plumes and ridges. *Earth Planet. Sci. Lett.* **128**, 425–437 (1994).
- Turner, S. *et al.* U-series isotopes and destructive plate margin magma genesis in the Lesser Antilles. *Earth Planet. Sci. Lett.* **142**, 191–207 (1996).
- Elliott, T., Plank, T., Zindler, A., White, W. & Bourdon, B. Element transport from slab to volcanic front at the Mariana arc. *J. Geophys. Res.* **102**, 14991–15019 (1997).
- Chabaux, F. & Allegre, C. J. ²³⁸U–²³⁰Th–²²⁶Ra disequilibria in volcanics: a new insight into melting conditions. *Earth Planet. Sci. Lett.* **126**, 61–74 (1994).
- Beattie, P. Uranium–thorium disequilibria and partitioning on melting of garnet peridotite. *Nature* **363**, 63–65 (1993).

- Hawkesworth, C. J., Turner, S. P., McDermott, F., Peate, D. W. & van Calsteren, P. U–Th isotopes in arc magmas: implications for element transfer from the subducted crust. *Science* **276**, 551–555 (1997).
- Woodhead, J., Eggin, S. & Gamble, J. High field strength and transition element systematics in island arc and back-arc basin basalts: evidence for multi-phase melt extraction and a depleted mantle wedge. *Earth Planet. Sci. Lett.* **114**, 491–504 (1993).
- Hawkesworth, C. J., Gallagher, K., Hergt, J. M. & McDermott, F. Mantle and slab contributions in arc magmas. *Annu. Rev. Earth Planet. Sci.* **21**, 175–204 (1993).
- Gill, J. B. & Williams, R. W. Th isotope and U-series studies of subduction-related volcanic rocks. *Geochim. Cosmochim. Acta* **54**, 1427–1442 (1990).
- McDermott, F. & Hawkesworth, C. J. Th, Pb and Sr isotope variations in young island arc volcanics and oceanic sediments. *Earth Planet. Sci. Lett.* **104**, 1–15 (1991).
- Condomines, M. & Sigmarsson, O. Why are so many arc magmas close to ²³⁸U–²³⁰Th radioactive equilibrium? *Geochim. Cosmochim. Acta* **57**, 4491–4497 (1993).
- Dupont, J. & Herzer, R. H. in *Geology and Offshore Resources of Pacific Island arcs—Tonga Region* (eds Scholl, D. W. & Vallier, T. L.) 323–332 (Circum-Pacific Council for Energy and Mineral Resources, Houston, 1985).
- Bevis, M. *et al.* Geodetic observations of very rapid convergence and back-arc extension at the Tonga arc. *Nature* **374**, 249–251 (1995).
- Ewart, A. & Hawkesworth, C. J. The Pleistocene–Recent Tonga–Kermadec arc lavas: Interpretation of new isotopic and rare earth data in terms of a depleted mantle source model. *J. Petrol.* **28**, 495–530 (1987).
- Cole, J. W., Graham, I. J. & Gibson, I. L. Magmatic evolution of Late Cenozoic volcanic rocks of the Lau Ridge, Fiji. *Contrib. Mineral. Petrol.* **104**, 549–554 (1990).
- Ewart, A., Bryan, W. B., Chappell, B. W. & Rudnick, R. L. Regional Geochemistry of the Lau–Tonga arc and backarc systems. *Proc. ODP Sci. Res.* **135**, 385–425 (1994).
- Hergt, J. M. & Hawkesworth, C. J. Pb-, Sr-, and Nd-isotopic evolution of the Lau Basin: implications for mantle dynamics during backarc opening. *Proc. ODP Sci. Res.* **135**, 505–517 (1994).
- Ewart, A., Brothers, R. N. & Meehan, A. An outline of the geology and geochemistry, and the possible petrogenetic evolution of the volcanic rocks of the Tonga–Kermadec–New Zealand island arc. *J. Volcanol. Geotherm. Res.* **2**, 205–250 (1977).
- Isacks, B. L. & Barazangi, M. in *Island Arcs, Deep Sea Trenches, and Back-Arc Basins* (eds Talwani, M. & Pitman, W. C.) 99–114 (Am. Geophys. Un., Washington DC, 1977).
- Turner, S. *et al.* ²³⁸U–²³⁰Th disequilibria, magma petrogenesis and flux rates beneath the depleted Tonga–Kermadec island arc. *Geochim. Cosmochim. Acta* (in the press).
- Hawkesworth, C., Turner, S., Peate, D., McDermott, F. & van Calsteren, P. Elemental U and Th variations in island arc rocks: implications for U-series isotopes. *Chem. Geol.* **139**, 207–221 (1997).
- Plank, T. & Langmuir, C. H. Tracing trace elements from sediment input to volcanic output at subduction zones. *Nature* **362**, 739–743 (1993).
- Miller, D. M., Goldstein, S. L. & Langmuir, C. H. Cerium/lead and lead isotope ratios in arc magmas and the enrichment of lead in the continents. *Nature* **368**, 514–520 (1994).
- Brenan, J. M., Shaw, H. F. & Ryerson, J. Experimental evidence for the origin of lead enrichment in convergent-margin magmas. *Nature* **378**, 54–56 (1995).
- Bourdon, B., Zindler, A., Elliott, T. & Langmuir, C. H. Constraints on mantle melting at mid-ocean ridges from global ²³⁸U–²³⁰Th disequilibrium data. *Nature* **384**, 231–235 (1996).
- Plank, T. & Langmuir, C. H. The geochemical composition of subducting sediment and its consequences for the crust and mantle. *Chem. Geol.* (in the press).
- Burns, R. E. *et al.* Site 204. *Init. Rep. DSDP* **21**, 33–56 (1973).
- Cheng, Q. *et al.* Isotopic evidence for a hotspot origin of the Louisville seamount chain. *Am. Geophys. Un. Monogr.*, Washington DC **43**, 283–296 (1987).
- Lonsdale, P. Sedimentation and tectonic modification of Samoan archipelagic apron. *Am. Assoc. Petrol. Geol. Bull.* **59**, 780–798 (1975).
- Peacock, S. M., Rushmer, T. & Thompson, A. B. Partial melting of subducting oceanic crust. *Earth Planet. Sci. Lett.* **121**, 227–244 (1994).
- Taylor, B., Zellmer, K., Martinez, F. & Goodliffe, A. Sea-floor spreading in the Lau back-arc basin. *Earth Planet. Sci. Lett.* **144**, 35–40 (1996).
- Dunn, T. & Sen, C. Mineral/matrix partition coefficients for orthopyroxene, plagioclase, and olivine in basaltic to andesitic systems: a combined analytical and experimental study. *Geochim. Cosmochim. Acta* **58**, 717–733 (1994).
- Spiegelman, M. & Elliott, T. Consequences of melt transport for uranium series disequilibrium in young lavas. *Earth Planet. Sci. Lett.* **118**, 1–20 (1993).
- Furukawa, Y. Magmatic processes under arcs and formation of the volcanic front. *J. Geophys. Res.* **98**, 8309–8319 (1993).
- McKenzie, D. The extraction of magma from the crust and mantle. *Earth Planet. Sci. Lett.* **74**, 81–91 (1985).
- Liu, J., Bohlen, S. R. & Ernst, W. G. Stability of hydrous phases in subducting oceanic crust. *Earth Planet. Sci. Lett.* **143**, 161–171 (1996).
- Green, D. H. Experimental melting studies on a model upper mantle composition at high pressure under water-saturated and water-undersaturated conditions. *Tectonophysics* **17**, 285–297 (1973).
- Wyllie, P. J. Magmas and volatiles components. *Am. Mineral.* **64**, 469–500 (1979).
- Davies, J. H. & Rowland, A. Importance of temperature dependent viscosity and hydraulic fracture on physical models of subduction zone magmatism. *Geol. Soc. Aust. Abstr.* **45**, 17–20 (1997).
- Nichols, G. T., Wyllie, P. J. & Stern, C. R. Subduction zone melting of pelagic sediments constrained by melting experiments. *Nature* **371**, 785–788 (1994).
- Molnar, P. & England, P. Temperatures in zones of steady-state underthrusting of young oceanic lithosphere. *Earth Planet. Sci. Lett.* **131**, 57–70 (1995).
- Danyushevsky, L. V., Sobolev, A. V. & Falloon, T. J. North Tongan high-Ca boninite petrogenesis: the role of Samoan plume and subduction zone-transform fault transition. *J. Geodyn.* **20**, 219–241 (1995).
- Palacz, Z. A. & Saunders, A. Coupled trace element and isotope enrichment in the Cook–Austral–Samoa islands, southwest Pacific. *Earth Planet. Sci. Lett.* **79**, 270–280 (1986).
- Wright, E. & White, W. M. The origin of Samoa: new evidence from Sr, Nd, and Pb isotopes. *Earth Planet. Sci. Lett.* **81**, 151–162 (1986).
- Jenner, G. A. *et al.* Determination of partition coefficients from trace elements in high pressure–temperature experimental run products by laser ablation microprobe–inductively coupled plasma–mass spectrometry (LAM–ICP–MS). *Geochim. Cosmochim. Acta* **58**, 5099–5103 (1994).

Acknowledgements. We thank H. Davies, T. Plank, D. Peate, N. Rogers, J. Hergt, I. Smith, T. Worthington, A. Cohen and P. van Calsteren for discussions about island arc magma petrogenesis and U-series disequilibria generally; J. Bartlett and N. Rogers for providing analytical assistance; I. Smith, T. Worthington, J. Pearce, S. Acland, T. Vallier and J. Gill for providing us with samples; T. Plank for allowing us to refer to her work in press; and T. Plank for reviews of the manuscript. This work was supported in part by the NERC, and S.T. was supported by a Royal Society University Research Fellowship.

Correspondence and requests for materials should be addressed to S.T. (e-mail: s.p.turner@open.ac.uk).

Competition between randomizing impacts and inelastic collisions in granular pattern formation

Troy Shinbrot

Department of Chemical Engineering, Northwestern University, Evanston, Illinois 60208, USA

The flow and mixing of granular materials occur during handling of a wide variety of substances, from pharmaceuticals to cement to cereal grains^{1,2}. The understanding of such flows is, however, considerably more limited than it is for fluids³; even basic processes such as tumbling^{4,5}, simple shear^{6–8} and shaking^{9–12} give rise to unexpected results. A case in point is granular pattern formation. A rich variety of patterns, including stripes, squares, hexagons and solitary structures, has been observed in vertically shaken, shallow granular beds^{13,14}. The vertical dynamics responsible for these patterns have been explored^{15–19}, but the role of horizontal motions of the grains is less well understood. Here I present a model of these motions that identifies two aspects as central to pattern formation: the randomization of horizontal velocities²⁰ by shaking, and the inelastic nature of grain collisions. These two elements alone, even without the influence of gravity, are sufficient to produce organized patterns in the horizontal plane—both those observed and others not yet seen experimentally.

High-speed photographs²¹ of granular patterns reveal two distinct stages of motion. First, the supporting plate accelerates upward to strike the granular bed. Patterns are at their sharpest at the moment of impact, and scatter apart thereafter. Second, the supporting plate accelerates downward faster than gravity (accelerations of 3–8g are typical). The granular bed then separates from the supporting plate, and particles collide to reform a patterned state, typically with the vertices of the new pattern at the centres of the previous pattern²².

Thus we examine a simplified model consisting of two sequential stages: once in every driving cycle particles scatter apart; thereafter they suffer inelastic collisions. The pattern formation mechanism in this model is straightforward. Consider for example two clusters of

particles separated by a distance λ , which scatter apart (in a way made precise below) as shown in Fig. 1a, b. Once particles from opposing clusters meet, they can collide along the line of symmetry¹² (dotted line), forming a column of nearly stationary particles as shown in Fig. 1c, d. Similar principles apply for more complicated symmetries. For example, in Fig. 1e–h, a square pattern emerges spontaneously from a random initial state under the influence of a simulation defined next.

We begin by examining the first, expanding, stage of motion. Two things of importance occur during this stage: particles scatter apart, and the bed dilates. Patterns are seen in granular beds ranging from 3 to 50 particles deep. In this situation, it seems reasonable to take the reaction of a stack of numerous convex particles to an energetic impact to be essentially random in direction^{23–26}. We explore this Ansatz by randomizing all particle velocities once per cycle according to:

$$U_a(i) = fU_b(i) + VU_r(i) \quad (1)$$

where $U_{a[b]}(i)$ is the velocity of the i th particle after [before] the impact, V and f are constants, and $U_r(i)$ is random in direction for each particle and is fixed in magnitude. The terms V and f respectively determine the strength of the impact and the particles' memories of their prior velocities. In the following results we choose $f = 0.5$.

The model is strictly two-dimensional and can be thought of as a horizontal projection of the true three-dimensional problem. To mimic dilation in this projection, we permit all particles to travel without collisions for a brief period, T_D , after the impact. Separate simulations reveal that the value of T_D is not crucial, but as $T_D \rightarrow 0$, particle motion becomes severely frustrated. We choose T_D to be one-fifth of the period between impacts, T .

After this dilation period, particles collide inelastically. This is the second stage of motion. We use a mean-field approach to compute collision effects. For each particle i we identify all neighbours, j : $1 \dots n_i$, that overlap with the i th particle. We then average the momenta of the n_i neighbouring particles and compute the reaction that the i th particle would undergo in a direct (that is, zero impact parameter) inelastic collision with a single particle having momentum given by the averaged value. We repeat this operation with every particle taking the role of the i th particle, and update particle velocities after all reactions have been computed. Collisions are non-binary and follow the rule

$$U_i \rightarrow \frac{(1 + n_i\epsilon)U_i + (1 - \epsilon)n_iU_j}{1 + n_i} \quad (2)$$

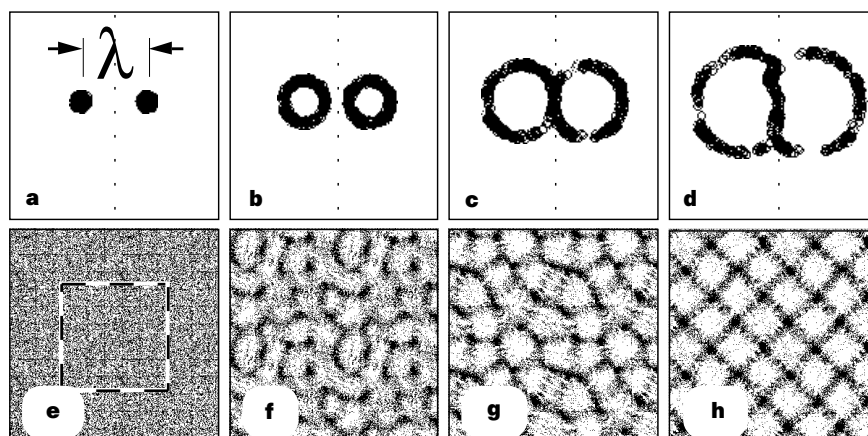


Figure 1 Evolution of particles subject to random expansion followed by inelastic collisions. **a, b**, Expansion of two clusters of 500 particles each; particles in each cluster are given an initial velocity random in direction and fixed in magnitude. **c, d**, Inelastic collisions between clustered particles occurs along symmetry line (dotted). **e–h**, Spontaneous emergence of patterned state from initially random

state. **e**, At time = 0, 32,767 identical particles are placed with zero velocity and random initial position. **f–h**, Snapshots of particle positions are shown respectively after 5, 10 and 100 simulated shakes of identical period T . Boundary conditions are periodic, and actual computation area (broken square in **e**) is duplicated to aid the eye.

where U_i and U_j are respectively the velocities of the i th particle and the mean velocity of its n_i neighbours, and ϵ is the effective coefficient of restitution for the collision. U_j is a mean field term and consequently individual collisions need not respect momentum conservation exactly. Nevertheless, every particle interacts with its mean field and momentum is conserved as an ensemble average. In the simulations below I use $\epsilon = 0.25$, and integrate trajectories with a time step $\Delta t = 5 \times 10^{-6}$ natural units. The patterned states expressed change imperceptibly as either ϵ or Δt is varied by 25%. Larger variations cause qualitative changes in particular patterns, but patterns still appear provided that $\epsilon \Delta t \ll T$.

Figure 2 shows results from multiple simulations, each using 32,767 particles. All snapshots are taken at the conclusion of 100 simulated shakes starting from random initial particle locations and zero initial velocities. In this figure the natural parameters, V and T , for the model are displayed. Experimental data are typically taken at fixed acceleration, so for comparison the figure also displays dispersion results keeping V/T constant. In the inset to Fig. 2 the wavelength, λ , of the observed pattern is plotted against the driving period, T , for simulations where $V/T \equiv 2 \times 10^7$ natural units (a value that travels through the centre of the main figure).

Wavelengths in the dispersion plot were obtained by autocorrelation analysis, and are best fit with the power law $\lambda \propto T^{1.9 \pm 0.1}$. This differs from experimental data^{21,27}, which give $\lambda \approx T^2$ for large periods, crossing over to $\lambda \approx T$ for smaller periods. Studies^{17,21} indicate that the crossover might be linked to a qualitative change in particle motion from lateral (which is captured in a horizontal projection) to vertical (which is not). My results support this conclusion to the extent that it is evidently possible to produce patterns in a horizontal plane in the absence of gravity or vertical motion. These patterns differ, however, from the physical, three-dimensional, situation in quantitative detail; moreover, patterns

have also been found in two-dimensional simulations confined to a vertical plane^{15,16}. These results taken together suggest that neither vertical nor horizontal simulations completely describe granular pattern formation. The problem can best be viewed as a competition between coupled processes acting in orthogonal directions^{27,28}.

Two new patterns are evident in Fig. 2. First, a triangular pattern is seen in the upper right corner of the figure. Because these patterns appear at high velocities and long periods, they may be beyond the capacity of standard laboratory shakers. Nevertheless, they are readily seen in these simulations and it may be possible to duplicate them experimentally. A second new pattern is seen as small striations within the hexagon at $V = 2,000$, $T = 2 \times 10^{-4}$, or the adjacent square at $T = 1.75 \times 10^{-4}$. The origin of these striations is not clear, but they seem to be most prevalent at longer periods. It would be interesting to know whether striations are found experimentally as well.

Two dimensionless groups seem to be most important in governing pattern selection. The first is the ratio $\Lambda = VT/d$, where d is the particle diameter. Pattern onset requires that $\Lambda \gg 1$, and in my simulations I find that Λ at onset ranges from 7 to 30, depending on parameter values. By comparison, experiments¹⁴ indicate a small wavelength cut-off ranging from five to eleven particle diameters. In Fig. 2 I include onset data (open circles), identified when peaks in autocorrelations of the particle positions first rise above the noise floor. The onset curve shown is the best fit to the power law $V \propto T^{-A}$, where $A = 0.8 \pm 0.2$.

The second dimensionless group is the effective Froude number⁹, $F = \langle U^2 \rangle / (Vd/T)$. This number characterizes dissipation, and $F = 1$ demarcates the boundary between high-dissipation patterns, for example stripes and irregular cells, and low-dissipation patterns, for example squares and triangles. The curve $F = 1$ shown is the best fit of the power law $V \propto T^{-B}$ where $B = 0.59 \pm 0.01$, to $\langle U^2 \rangle$ versus

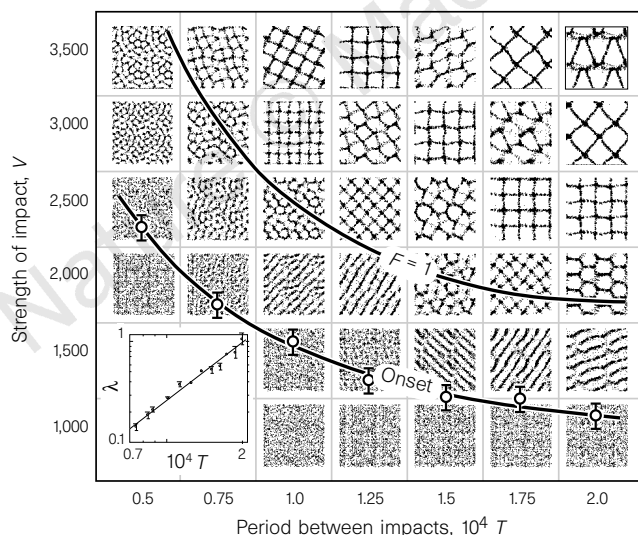


Figure 2 Patterned states as a function of the strength of the randomizing tap (V) and the period between taps (T). The regular patterns (stripes, squares, hexagons and triangles) repeat every other cycle, whereas the irregular patterns remaining change spatiotemporally. No patterns are seen below the onset curve indicated; data for this curve were obtained using autocorrelation analysis (see text). The boundary between high-dissipation patterns (for example, stripes and irregular cells) and low-dissipation patterns (for example, squares and triangles) occurs at Froude number $F = 1$. Periodic boundary conditions are used; four copies of the computational domain (as defined in Fig. 1e) are displayed to aid the eye. Inset: wavelength of patterns plotted against driving period for simulations holding V/T constant. The particle radius is reduced from 0.01 units (main plot) to 0.00544 units (inset) to extend the small-wavelength cut-off. The power-law fit suggests a nearly quadratic dispersion relation.

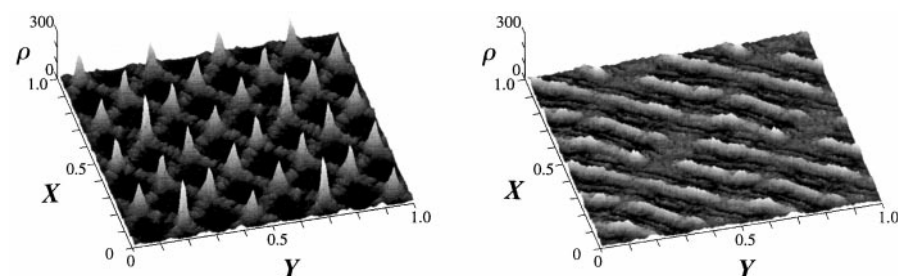


Figure 3 Particle densities from mean-field simulation in square (left), and striped (right) patterns. Densities are obtained from particle positions shown in Fig. 2, ($V = 2,500$, $T = 1.25 \times 10^{-4}$) and ($V = 1,500$, $T = 1.75 \times 10^{-4}$) respectively. Densities are sharply clustered in the square state and are more diffuse in the striped state.

T data obtained from separate simulations. The source of the exponents A and B is not clear, but implications can be drawn from their relative magnitudes. As $A > B$, the separation between the onset and the curve $F = 1$ should diminish as the driving period grows. This means that for long periods, stripes should give way to square or other regular cellular patterns, whereas for shorter periods, irregular structures should dominate.

The significance of dissipation on pattern selection has been commented on previously^{9,21}. For example, it has been observed that stripes are present only in high-density, high-dissipation experiments; similar results are obtained in fluids experiments^{29,30}. My simulations are consistent with these results: I find that the transition between stripes and square moves closer to onset and then vanishes altogether, as the density, ρ , is decreased below $\rho \approx 3$ (measured in units of total particle area divided by available area).

In addition, we see very different densities and dissipation rates for patterned states on opposite sides of the curve $F = 1$. Figure 3 shows representative densities for nearby striped and square patterns. Particles in the square pattern are much more strongly clustered around the vertices, with maximum densities an order of magnitude larger than in the striped pattern. The energies dissipated are very different as well: the mean kinetic energy at the conclusion of a cycle is 270% higher for the square than for the striped pattern.

In conclusion, it may on the one hand seem surprising that an exceedingly simplified model, containing little more than a combination of periodic randomization and dissipation, and entirely neglecting gravity, can produce an apparent wealth of spontaneously organized and highly structured patterns. On the other hand, at a fundamental level there is little difference between this construction and well-known reaction–diffusion models^{31,32}. There, too, randomization (in the form of diffusion) competes with dissipation (in the form of reactions) to produce a rich tapestry of regular and irregular patterns. Future experiments may clarify connections between granular and other pattern formation mechanisms. □

Received 18 April; accepted 22 July 1997.

- Jaeger, H. M., Nagel, S. R. & Behringer, R. P. Granular solids, liquids and gases. *Rev. Mod. Phys.* **68**, 1259–1273 (1996).
- Ennis, B. J., Green, J. & Davies, R. The legacy of neglect in the US. *Chem. Eng. Progr.* **90**, 32–43 (1994).
- Ottino, J. M. & Shinbrot, T. Comparing extremes: mixing of fluids, mixing of solids. *NATO-ASI Mixing: Chaos and Turbulence Proc.* (in the press).
- Metcalfe, G., Shinbrot, T., McCarthy, J. J. & Ottino, J. M. Avalanche mixing of granular solids. *Nature* **374**, 39–41 (1995).
- Hill, K. M. & Kakaliotis, J. Reversible axial segregation of binary mixtures of granular materials. *Phys. Rev. E* **49**, R3610–R3614 (1994).
- Bagnold, R. A. The shearing and dilatation of dry sand and the 'singing' mechanism. *Proc. R. Soc. Lond. A* **295**, 219–232 (1966).
- Nasuno, S., Kudrolli, A. & Gollub, J. P. Friction in granular layers: hysteresis and precursors. *Phys. Rev. Lett.* **79**, 944–952 (1997).
- Miller, B., O'Hern, C. & Behringer, R. P. Stress fluctuations for continuously sheared granular materials. *Phys. Rev. Lett.* **77**, 3110–3113 (1996).
- Umbanhowar, P. B., Melo, F. & Swinney, H. L. Localized excitations in a vertically vibrated granular layer. *Nature* **382**, 793–796 (1996).
- Luding, S., Clément, E., Blumen, A., Rajchenbach, J. & Duran, J. Studies of columns of beads under external vibrations. *Phys. Rev. E* **49**, 1634–1610 (1994).
- Clément, E., Vanel, L., Rajchenbach, J. & Duran, J. Pattern formation in a vibrated granular layer. *Phys. Rev. E* **53**, 2972–2979 (1996).
- Pak, H., van Doorn, E. & Behringer, R. Effects of ambient gases on granular materials under vertical vibration. *Phys. Rev. Lett.* **74**, 4643–4647 (1995).
- Metcalfe, T. H., Knight, J. B. & Jaeger, H. M. Standing wave patterns in shallow beds of vibrated granular material. *Physica A* **236**, 202–210 (1997).
- Melo, F., Umbanhowar, P. & Swinney, H. L. Transition to parametric wave patterns in a vertically oscillated granular layer. *Phys. Rev. Lett.* **72**, 172–175 (1994).
- Doudady, S., Fauve, S. & Laroche, C. Subharmonic instabilities and defects in a granular layer under vertical vibrations. *Europhys. Lett.* **8**, 621–627 (1989).
- Wassgren, C. R., Brennen, C. E. & Hunt, M. L. Vertical vibration of a deep bed of granular material in a container. *J. Appl. Mech.* **63**, 712–719 (1996).
- Melo, F., Umbanhowar, P. B. & Swinney, H. L. Hexagons, kinks, and disorder in oscillated granular layers. *Phys. Rev. Lett.* **75**, 3838–3841 (1995).
- Cerda, E., Melo, F. & Rica, S. A model for parametric waves in granular materials (preprint). (Dept. de Física, Univ. de Santiago, 1997).
- Luding, S., Clément, E., Rajchenbach, J. & Duran, J. Simulations of pattern formation in vibrated granular media. *Europhys. Lett.* **36**, 247–251 (1996).
- Menon, N. & Durian, D. J. Diffusing-wave spectroscopy of dynamics in a three-dimensional granular flow. *Science* **275**, 1920–1922 (1997).
- Metcalfe, T. H., Knight, J. B. & Jaeger, H. M. Standing wave patterns in shallow beds of vibrated granular material. *Physica A* **236**, 202–210 (1997).

- Crawford, J. D., Gollub, J. P. & Lane, D. Hidden symmetries of parametrically forced waves. *Nonlinearity* **6**, 119–164 (1993).
- Sinai, Ya. G. Dynamical systems with elastic reflections: ergodic properties of dispersing billiards. *Russ. Math. Surv.* **25**, 137–147 (1970).
- Bleher, S., Grebogi, C. & Ott, E. Bifurcation to chaotic scattering. *Physica D* **46**, 87–121 (1990).
- Chen, Q., Ding, M.-Z. & Ott, E. Chaotic scattering in several dimensions. *Phys. Lett. A* **115**, 93–95 (1990).
- Ding, M.-Z., Grebogi, C., Ott, E. & Yorke, J. A. Transition to chaotic scattering. *Phys. Rev. A* **42**, 7025–7040 (1990).
- Bizon, C., Shattuck, M. D., Swift, J. B., McCormick, W. D. & Swinney, H. L. Patterns in 3D vertically oscillated granular layers: simulation and experiment. *Phys. Rev. Lett.* (in the press).
- Tsimring, L. & Aranson, I. Localized and cellular patterns in a vibrated granular layer. *Phys. Rev. Lett.* **79**, 213–216 (1997).
- Kudrolli, A. & Gollub, J. P. Localized spatiotemporal chaos in surface waves. *Phys. Rev. E* **54**, R1052–R1055 (1996).
- Edwards, W. & Fauve, S. Patterns and quasi-patterns in the Faraday experiment. *J. Fluid Mech.* **278**, 123–134 (1994).
- Winfrey, A. T. Spiral waves of chemical activity. *Science* **175**, 634–638 (1972).
- Pearson, J. E. Complex patterns in a simple system. *Science* **261**, 189–192 (1993).

Acknowledgements. I thank J. M. Ottino for thoughtful advice, J. J. McCarthy for valuable comments and assistance, and H. Jaeger, M. Mavrouniotis, H. Swinney and P. Umbanhowar for crucial scientific discussions. This work was supported by the National Science Foundation (Division of Chemical and Transport Systems) and the American Chemical Society (Petroleum Research Fund).

Correspondence and requests for materials should be addressed to the author (e-mail: shinbrot@nwu.edu).

Induced long-range order in crosslinked 'one-dimensional' stacks of fluid monolayers

Gerald C. L. Wong*, Wim H. de Jeu*, Henry Shao†, Keng S. Liang‡ & Rudolf Zentel‡

* FOM Institute for Atomic and Molecular Physics, Kruislaan 407, 1098SJ Amsterdam, The Netherlands

† Exxon Research and Engineering Company, Annandale, New Jersey 08801, USA

‡ Fachbereich 9: Chemie und Biologie, Bergische Universität Wuppertal, Gauss-Strasse 20, D-42097 Wuppertal, Germany

Ordinary crystals are characterized by long-range translational order in all three dimensions. In lower-dimensional systems, in contrast, translational order is destroyed through the 'Landau–Peierls instability'—displacements from periodic ordering due to thermal fluctuations whose amplitude increases with the size of the system^{1–4}. This effect is well known for layered systems ordered in one dimension, such as surfactant membranes^{5,6}, smectic (layered) liquid crystals⁷ and liquid crystalline polymers⁸, which form ordered stacks of fluid monolayers. Smectic liquid-crystal polymers can be weakly crosslinked to form percolating elastomeric networks that still allow mobility on a molecular scale^{9,10}. In these smectic elastomers, fluctuations of the fluid layers are coupled to distortions of the underlying network, and are therefore energetically penalized¹¹, even though the network of crosslinks has a random nature and thus no three-dimensional translational order. Here we present a high-resolution X-ray diffraction study of a smectic elastomer that reveals the effects of crosslinking on long-range ordering. We find that the introduction of a random network of crosslinks enhances the stability of the layered structure against thermal fluctuations and suppresses the Landau–Peierls instability so as to induce 'one-dimensional' long-range ordering at length-scales up to several micrometres.

Landau¹ and Peierls² were the first to demonstrate that translational order is destroyed in one- and two-dimensional systems by thermal fluctuations. In three-dimensional space similar arguments can be applied to, for example, a smectic system of stacked fluid monolayers, where rod-like mesogenic (liquid-crystal-forming) molecules order into a 'one-dimensional' density wave along one direction, but remain fluid in the other two (Fig. 1a). In such a system the mean squared layer displacement diverges logarith-

mically with the system size at all finite temperatures ('Landau–Peierls instability'). The resultant X-ray diffraction signature has been calculated by Caillé¹². In the direction perpendicular to the layers, these divergent thermal fluctuations transform the discrete Bragg peaks (indicative of translational order) into algebraic 'cusps' with the structure factor S having an asymptotic power-law form:

$$S(0, 0, q_z) \approx \frac{1}{|q_z - q_m|^{2-\eta_m}}$$

where q_z represents the wave vector transfer along the density wave, q_m represents the position of the m th order of diffraction ($q_m = mq_0$, $q_0 = 2\pi/d$, $m = 1, 2, 3, \dots$, and d is the smectic layer spacing). The exponent η_m is given by $\eta_m = q_m^2 k_B T / (8\pi(BK)^{1/2})$ and directly related to the elastic constants of this smectic system: K is the bending modulus and B is the layer compression modulus (T is temperature, and k_B is Boltzmann's constant). The magnitude of η_m is quite small, with typical measured values of ~ 0.1 – 0.3 near the smectic-A/nematic phase transition temperature (T_{AN}), where B approaches zero. At all other temperatures, $\eta_m \approx 0$, which results in a saturated lineshape with 'tails' given by $1/q^2$. This can make the discrimination between Caillé lineshapes and normal Bragg peaks difficult, as Bragg peaks broaden into an asymptotic $1/q^2$ form as well, due to contributions from dynamical effects and thermal diffuse scattering¹³.

In a smectic polymer (Fig. 1b) the mesogenic molecules are attached to polymer chains via flexible spacer groups. Although the phase transition temperatures can differ significantly from those of its monomeric counterpart due to the coupling between the 'one-dimensional' layering and the three-dimensional ensemble of polymer chains, the Landau–Peierls instability persists in such smectic polymers: the Bragg peaks are still destroyed by fluctuations⁸. These liquid-crystalline polymers can be weakly crosslinked into an elastomer network (Fig. 1c) with almost no change in the phase transition temperatures. The macroscopic rubber elasticity introduced via such a network¹⁴ interacts with the liquid crystalline ordering. In the case of nematic elastomer networks, where the mesogenic units are aligned but not layered, novel forms of mechanical instabilities and orientational memory effects have been observed¹⁵. In a smectic elastomer, the fluctuating smectic layers cannot move past crosslinks in the network: this network coupling imposes a penalty for relative translations between the orientationally ordered fluid layers and the polymer network. A recent continuum theory has predicted that the fluctuations associated with the Landau–Peierls instability can be suppressed in a smectic elastomer, so that the usual logarithmic divergence of the mean squared layer displacements no longer exists¹¹. The marginally stable, quasi-periodic structure of normal 'one-dimensional' fluid stacks should be transformed into a fully periodic structure with proper long-range order, and the resultant diffraction should change from Caillé lineshapes to normal Bragg peaks.

A smectic elastomer system has a significant degree of random

disorder, which is not included in the original theoretical model (ref. 11). When such a random field is taken into account the quenched static disorder will eventually destroy translational long-range order at large length-scales (small q values), so that the discrimination between Bragg peaks and Caillé lineshapes must be made at large enough q values. We note that the Caillé lineshape cannot have an asymptotic q -dependence that is steeper than $1/q^2$, whereas the behaviour of a true Bragg peak is not bound by this constraint. Even a Bragg peak that has been broadened by lattice imperfections¹⁶ can in principle still be sharper than $1/q^2$, and may not reach this asymptotic limit for several decades of intensity.

The basic macromolecular system used in this study is a smectic random copolymer with a comb-like architecture¹⁷. Most of the side chains connect rigid 'rod-like' molecules to the polyacrylate backbone, which provide the system with liquid crystalline phase behaviour (Fig. 2). A small fraction ($\sim 5\%$) of the side-chains, however, terminate with functional hydroxyl groups, and serve as active sites for the crosslinking process. These 'linear' liquid-crystalline copolymer chains can be transformed into a weakly linked, 'infinite-molecular-weight' network via reaction with a crosslinking agent in a solvent (Fig. 1c and Fig. 2). The measurements reported below were performed at beamline X-10A of the National Synchrotron Light Source, Brookhaven National Laboratory, USA. To get monodomain samples, the uncrosslinked polymer is aligned with a magnetic field in a temperature-controlled cell. In contrast, the liquid-crystalline groups in the crosslinked elastomer are aligned with a strain field, which is applied *in situ* to freely suspended samples in a different temperature-controlled cell with an internal translation stage. Owing to the multiple-Bragg reflections in the monochromator and the analyser crystals, the measured intensity of the incident beam has an approximately $1/q^{3.6}$ dependence, which is much sharper than $1/q^2$ and therefore sufficiently steep for the present lineshape measurement. The in-plane liquid-like disorder has been independently confirmed by wide-angle X-ray scattering.

Measurements for both the uncrosslinked smectic polymer and the crosslinked smectic elastomer were done at $T = 55.0^\circ$, which is in the smectic-A phase well below T_{AN} and well above the glass transition temperature T_g . The full mosaic width, which measures the variation of layering directions about the average, is 2.5° for the elastomer and 4.1° for the polymer. Figure 3a shows the scattering intensity from the first-order diffraction of the aligned elastomer. The scattering intensity for the smectic elastomer decreases rapidly away from q_1 , with a slope of -2.40 ± 0.10 on a log–log plot over three orders of magnitude. The Bragg scattering from the elastomer is significantly sharper than a Caillé power-law lineshape, which saturates at a limiting slope of -2 at these temperatures. Figure 3b compares the asymptotic slopes of the diffuse scattering from the crosslinked elastomer and the corresponding uncrosslinked polymer. The dramatic sharpening of the elastomer lineshape is correlated with the existence of percolating crosslinks and the resultant change in polymer topology. In contrast to the lineshape of the

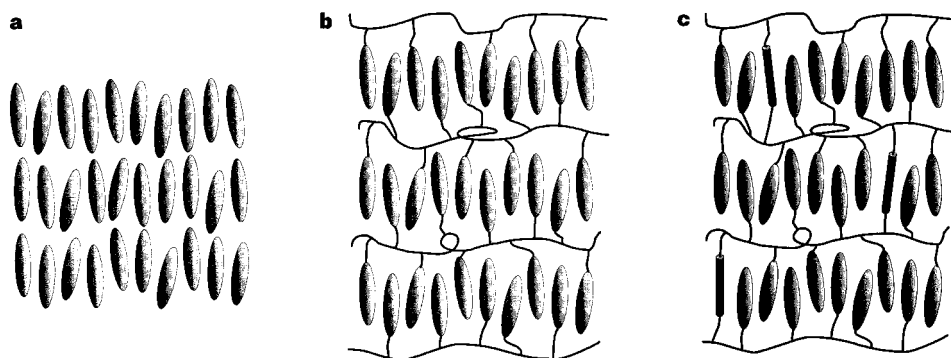


Figure 1 Schematic representation of a smectic liquid crystal (a), a smectic-crystalline polymer (b), and a weakly crosslinked smectic elastomer network (c). Crosslinks are represented by cylinders. The low density of crosslinks provides the smectic elastomer system with solid-like behaviour at macroscopic length-scales while retaining liquid-like behaviour at microscopic length-scales.

elastomer, the intensity tails of the uncrosslinked polymer have a slope of -1.85 ± 0.10 . The absolute value of the measured slope is expected to be somewhat less than the theoretical limit of 2 owing to the finite mosaic width of the sample. In fact, the mosaic distribution causes the Caillé lineshape to vary continuously between a power law with an exponent of $2 - \eta_m$ (for a single-orientation sample with no mosaic spread) to one with an exponent of $1 - \eta_m$ (for a random powder)¹⁸. Moreover, we have performed measurements on a sample of the same compound with a narrower mosaic width, which is possible at temperatures closer to T_{AN} , and obtained the established saturation value of -2 for the asymptotic slope.

We emphasize that the smectic elastomer network remains liquid at microscopic length-scales, even though it is a soft solid at macroscopic length-scales. Macroscopic deformations of the sample were found to drastically change the orientational distribution and layering texture at the molecular level. This is also evident in preliminary measurements of the weak higher-order diffraction (data not shown), which is sensitive to the 'sharpness' of layer interfaces at length-scales smaller than a smectic layer spacing, and is therefore susceptible to such liquid-like disorder. More importantly, the sample can melt into its nematic and isotropic phases on heating and reorder into the present rubber-like smectic phase. In other words, the smectic elastomer forms a soft, thermodynamically stable 'one-dimensional' lattice, and not a quenched 'glassy' solid with frozen-in order. In addition, the characteristic timescale that separates solid-like behaviour and liquid-like behaviour decreases as T_{AN} is approached from lower temperatures. Consequently the dynamical coupling responsible for the suppression of the Landau–Peierls instability is expected to disappear near T_{AN} (ref. 19).

Owing to its random nature, the elastomer lattice is expected to

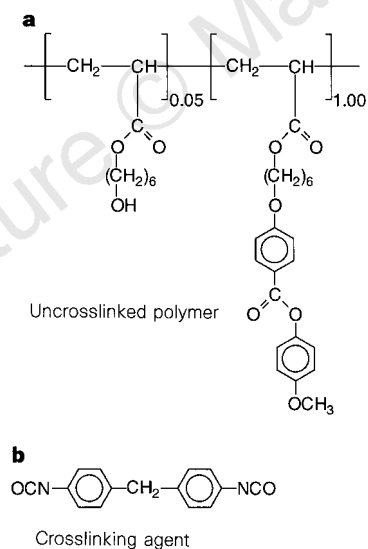


Figure 2 The polymer and crosslinking agent used in these experiments. **a**, The uncrosslinked polyacrylate-based, side-chain liquid-crystalline polymer (number-average relative molecular mass, 11,200), synthesized by radical copolymerization. The proportion of side-chains which terminate with functional hydroxyl groups is 5 mol.%. **b**, Crosslinking is achieved by reaction with 4,4'-diphenylmethane diisocyanate under basic conditions (triethylamine) in toluene. The similarity in the phase sequences of the elastomer (g 31 S_A 80 N 111 I) and the uncrosslinked polymer (g 26 S_A 82 N 110 I) has been verified with differential scanning calorimetry and polarizing microscopy (temperatures given in °C; g, S_A and N indicate the glassy, smectic-A and nematic phases, respectively). The crosslink density of the smectic elastomer network is ~2–5 mol.%, as estimated from a swelling experiment in toluene.

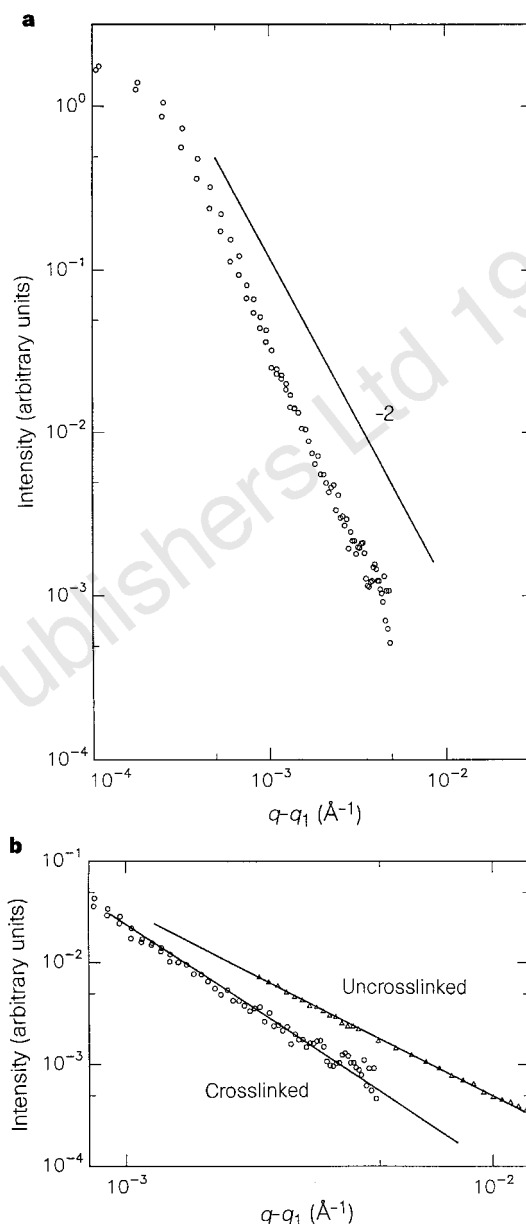


Figure 3 Results of X-ray diffraction studies. **a**, A log-log plot of the background-subtracted first-order diffraction intensity from the smectic elastomer network. The diffraction intensity decreases sharply away from q_1 and is significantly sharper than the normal Caillé lineshape for stacked fluid layers (limiting slope, -2). **b**, Comparison of the asymptotic diffraction intensity tails from the elastomer (circles; $q_1 = 0.255 \text{ \AA}^{-1}$, slope of 2.40 ± 0.10) and the corresponding uncrosslinked polymer (triangles; $q_1 = 0.253 \text{ \AA}^{-1}$, slope of 1.85 ± 0.10). The sharpening of the elastomer lineshape is clearly associated with the existence of crosslinks. The asymptotic slope is reached at slightly different q ranges for the two samples owing to domain size effects. The uncrosslinked polymer was sealed in a quartz capillary (1.5 mm diameter), and placed between the poles of a 0.8 T SmCo₅ magnet, inside the inner stage of a two-stage sample cell with a temperature stability of ± 30 mK. Alignment of the smectic planes in the uncrosslinked polymer was achieved by cooling the sample from the nematic phase at $T = 100^\circ\text{C}$ in the magnetic field, at a rate of 0.2°C h^{-1} . The elastomer sample ($\sim 7 \times 2 \times 1$ mm) was stretched *in situ* by 25% in the nematic phase at 90°C , then subsequently cooled into an aligned smectic phase. The sample cells were coupled to a 4-circle diffractometer, with q_z orientated perpendicular to the smectic layers. The monochromator and the analyser consisted of a double-bounce Si(111) and a triple-bounce Ge(111) channel-cut crystal set non-dispersively at a wavelength of $\lambda = 1.5347 \text{ \AA}$, which resulted in a sharp in-phase resolution of $\Delta q_z = 1.6 \times 10^{-4} \text{ \AA}^{-1}$, given by the half-width at half-maximum of the incident beam.

Origin of rubber-like behaviour in metal alloys

Xiaobing Ren & Kazuhiro Otsuka

Institute of Materials Science, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

have a high density of defects. It is interesting to compare these defects with flux vortex lattices in superconductivity²⁰. At large enough length-scales (above the so-called Larkin length), even arbitrarily weak static defects can destroy translational long-range order for all dimensions < 4 . For flux vortex lattices, the defects are rigidly fixed in space, and not connected to the fluctuating lattice. Long-range order is therefore always destroyed at large enough length-scales. The analogy with smectic elastomers is complex, because the defects are crosslinks which are embedded in the fluctuating lattice itself, and therefore are not strictly static. We observed $1/q^{2.4}$ intensity tails which persist down to q values approaching our experimental resolution limit, which is high enough to resolve length-scales of several micrometres. For the limiting case of static crosslinks, the scattering signature may consist of a crossover between the disordered Larkin limit at a small q , to the predicted Bragg behaviour at large q , which is the range we observe. As the Larkin length has been known to be macroscopically large even for soft lattices (tens of micrometres) we may not be able to observe this effect even at the high resolution of our experiments. But because crosslinks can locally disturb the smectic layering, and can even destroy the 'one-dimensional' lattice down to molecular length-scales at high crosslink densities²¹, the lineshape at our measured range of q values can also be affected. As with effects related to the mosaic width, such defects can only broaden rather than sharpen the lineshape associated with the 'one-dimensional' ordering, and therefore cannot explain our unambiguous observation of an elastomer Bragg peak that is sharper than the Caillé lineshape. In these smectic elastomers the introduction of disorder in the form of a random network has the counterintuitive result of enhancing the 'one-dimensional' ordering. □

Received 29 November 1996; accepted 15 August 1997.

- Landau, L. D. in *Collected Papers of L. D. Landau* (ed. ter Haar, D.) 193–216 (Gordon & Breach, New York, 1965).
- Peierls, R. E. Bemerkung über Umwandlungstemperaturen. *Helv. Phys. Acta Suppl.* **7**, S81–S83 (1934).
- Chaikin, P. M. & Lubensky, T. C. *Principles of Condensed Matter Physics* (Cambridge Univ. Press, New York, 1995).
- de Gennes, P. G. & Prost, J. *The Physics of Liquid Crystals* 2nd edn (Clarendon Press, Oxford, 1993).
- Safinya, C. R. *et al.* Steric interactions in a model multimembrane system: a synchrotron X-ray study. *Phys. Rev. Lett.* **57**, 2718–2721 (1986).
- Sirota, E., Smith, G. S., Safinya, C. R., Plano, R. J. & Clark, N. A. X-ray scattering studies of aligned, stacked surfactant membranes. *Science* **242**, 1406–1409 (1988).
- Als-Nielsen, J. *et al.* Observation of algebraic decay of positional order in a smectic liquid crystal. *Phys. Rev. B* **22**, 312–320 (1980).
- Nachaliel, E., Keller, E. N., Davidov, D. & Boeffel, C. Algebraic dependence of the structure factor and possible anharmonicity in a high resolution X-ray study of a side-group polymeric liquid crystal. *Phys. Rev. A* **43**, 2897–2902 (1991).
- Zentel, R. Liquid crystalline elastomers. *Angew. Chem. Adv. Mater.* **101**, 1407–1415 (1989).
- Finkelmann, H., Benne, J. & Semmler, K. Smectic liquid crystal elastomers. *Macromol. Symp.* **96**, 169–174 (1995).
- Terentjev, E. M., Warner, M. & Lubensky, T. C. Fluctuations and long-range order in smectic elastomers. *Europhys. Lett.* **30**, 343–348 (1995).
- Caillé, A. Remarques sur la diffusion des rayons X dans les smectiques. *C. R. Acad. Sci. Paris* **B274**, 891–893 (1972).
- Warren, B. E. *X-ray Diffraction* (Addison-Wesley, Reading, 1969).
- de Gennes, P. G. in *Polymer Liquid Crystals* (eds Ciferri, A., Krighbaum, W. R. & Meyer, R. B.) 115–132 (Academic, New York, 1982).
- Mitchell, G. R., Davis, F. J. & Guo, W. Strain-induced transitions in liquid-crystal elastomers. *Phys. Rev. Lett.* **71**, 2947–2950 (1993).
- Krivoglaz, M. *Theory of X-ray and Thermal Neutron Scattering by Real Crystals* (Plenum, New York, 1969).
- Zentel, R. & Benalia, M. Stress-induced orientation in lightly crosslinked liquid-crystalline side-group polymers. *Makromol. Chem.* **188**, 665–674 (1987).
- Kaganer, V. M., Ostrovsky, B. I. & de Jeu, W. H. X-ray scattering in smectic liquid crystals: from ideal-to real-structure effects. *Phys. Rev. A* **44**, 8158–8166 (1991).
- Olmsted, P. D. & Terentjev, E. M. Mean-field nematic-smectic-A transition in a random polymer network. *Phys. Rev. E* **53**, 2444–2453 (1996).
- Blatter, G., Feigelman, M. V., Geshkenbein, V. B., Larkin, A. I. & Vinokur, V. M. Vortices in high temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1380 (1994).
- Terentjev, E. M. Non-uniform deformations and molecular random fields in liquid crystalline elastomers. *Macromol. Symp.* **117**, 79–88 (1997).

Acknowledgements. We thank E. M. Terentjev for discussions, E. Gebhard for assistance in sample preparation, and J. Commandeur for technical support. This work is part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie [Foundation for Fundamental Research on Matter (FOM)] with support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek [Netherlands Organization for the Advancement of Research (NWO)]. W.d.J. acknowledges additional support from a NATO collaborative research grant.

Correspondence and requests for materials should be addressed to W.H.d.J. (e-mail: deju@amolf.nl)

Since 1932 it has been known that a number of ordered alloys show an unusual kind of deformation behaviour^{1–3}. These alloys (including Au–Cd, Au–Cu–Zn, Cu–Zn–Al, Cu–Al–Ni)^{4–8}, after being aged for some time in a martensitic state (the low-symmetry phase of a diffusionless transformation), can be deformed like a soft and pseudo-elastic rubber (with a recoverable strain as large as a few per cent). Accompanying martensite ageing is the development of martensite stabilization⁹ (increase in the temperature of reverse transformation to the parent state), the avoidance of which is important in actuator applications of the shape-memory effect²⁹, (which these alloys also generally exhibit). The origin of this rubber-like behaviour and of the ageing effect has remained unclear^{10–17}. Here we show that this behaviour does not involve a change in the degree of long-range order, but is instead due to an atomic rearrangement within the same sublattice of the imperfectly ordered alloy during martensite ageing. This process is driven by a general tendency for the equilibrium symmetry of the short-range order configuration of lattice imperfections to conform to the symmetry of the lattice. This principle not only explains all the observed aspects of the rubber-like behaviour and the ageing effect in both ordered and disordered alloys, but may also further our understanding of some diffusion phenomena in other crystalline materials.

A martensitic transformation lowers the symmetry of a crystal without involving atomic exchange or diffusion. As a result of symmetry lowering, a single crystal of the parent phase (high-temperature phase) is split into many twin-related martensite domains. If the martensite of the above-mentioned alloys is deformed immediately after being transformed from the parent phase, the strain can be accommodated by the easy reversal of some of the domains into new ones, that is, twinning. This explains the 'softness' of the martensite. Because the stress-induced domains have the same martensite structure and hence the same stability as

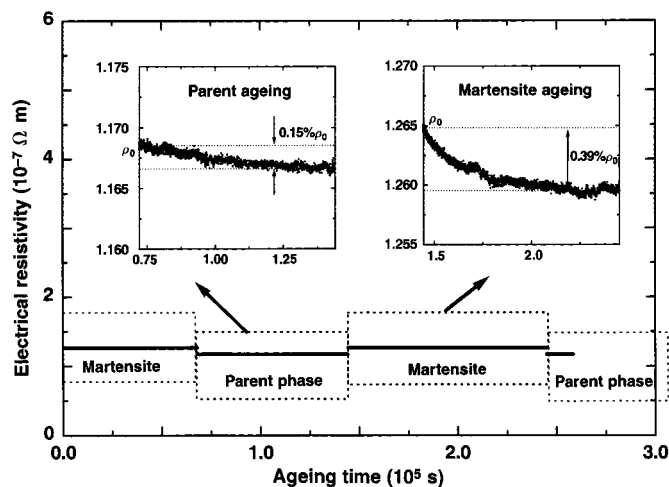


Figure 1 Electrical resistivity of single crystal Au_{50.5}Cd_{49.5} alloy as a function of ageing time, for the parent phase (enlarged in the left inset) after transformation from well-aged martensite, and for martensite (enlarged in the right inset) after transformation from well-aged parent.

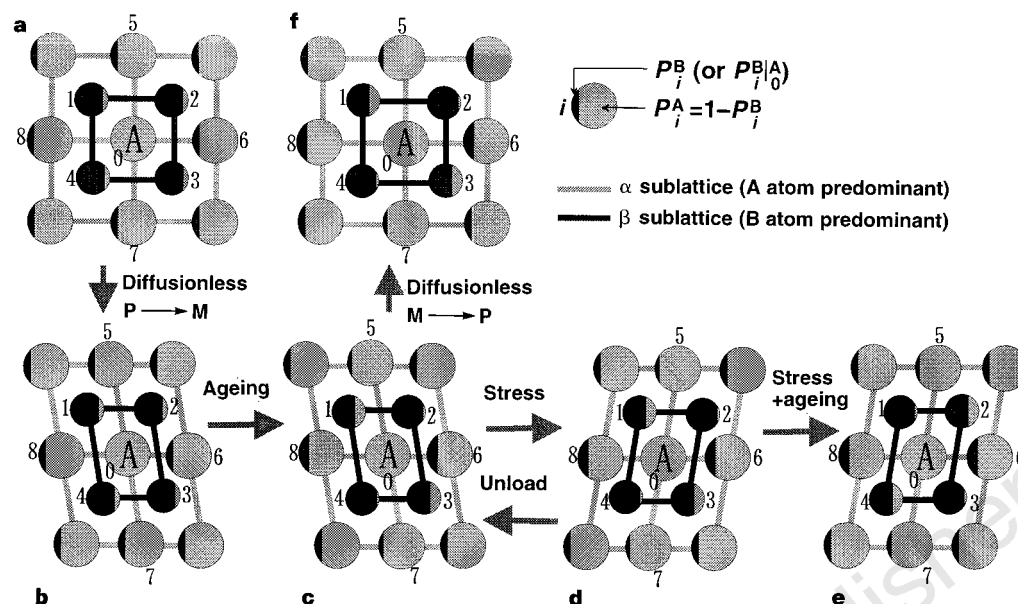


Figure 2 Mechanism of rubber-like behaviour and the martensite stabilization phenomenon. The illustrations show the statistical atomic configuration (conditional probabilities around an A atom) of an imperfectly-ordered A–B alloy in: **a**, the equilibrium parent phase; **b**, the martensite immediately after transformation from **a**; **c**, the equilibrium martensite; **d**, the stress-induced martensite domain (twin) formed immediately from **c**; **e**, the equilibrium state of the stress-induced domain; **f**, the parent transformed immediately from **c**. P, parent phase; M, martensite. P_i^B (or P_i^A) is the conditional probability of a B atom (or an A atom) occupying site i ($i = 1, 2, 3, \dots, 8$) if an A atom is at site 0. The relative values of P_i^B and P_i^A are indicated by the black and grey areas respectively.

the original ones, the original domains cannot be restored even when stress is released. This leads to a plastic or irrecoverable strain for the ‘fresh’ martensite.

However, the rubber-like behaviour appears when martensite is deformed after being aged for a certain period. Then the stress-induced domains seem to be less stable than the original ones, so the original domains can be restored when the stress is removed. This leads to a recoverable strain and to macroscopic rubber-like behaviour (RLB). The central problem is to explain what happens during martensite ageing to cause the stability difference between the stress-induced domain and the original one. RLB is different from two other kinds of pseudoelasticity of known origin. One is due to stress-induced martensitic transformation¹, and the other is due to the formation of a pseudo-twin¹⁸. In both cases, stress induces a structural transformation that provides the restoring force.

The existence of RLB in single-domain martensites^{19,20} proved that RLB and the ageing effect must originate from atomic rearrangement throughout the martensite. Previous martensite reconstruction models^{11,12} were found to contradict experimental observations²¹. Recent models^{13–17} also predicted some change in the average martensite structure (that is, the degree of long-range order (LRO)) during ageing. However, X-ray experiments on stable (that is, no tendency to decompose) Au–Cu–Zn²² and Au–Cd^{20,21} martensites did not provide any evidence for a change in the degree of LRO, but with a relatively large uncertainty (a few per cent).

Here we present experimental evidence to show that no change in LRO occurs during ageing. Because of the great sensitivity of electrical resistivity to changes in LRO in alloys, we used a four-terminal electrical resistivity measurement to monitor the possible LRO change during isothermal ageing for a Au_{50.5}Cd_{49.5} ordered alloy in a consecutive ageing sequence: martensite (298 ± 0.2 K for ~2 × 10⁵ s) → parent (330 ± 0.2 K for ~10⁵ s) → martensite (298 ± 0.2 K for ~10⁵ s). The result is shown in Fig. 1. Notably, both the parent and the martensite show a small decrease in resistivity (relaxation type) with ageing time before reaching saturation. Because diffusionless martensitic transformation does not change LRO, no change in LRO will occur during ageing if the parent and martensite have the same equilibrium LRO. If they were different, one would expect an increase in LRO in one phase and a decrease in the other during ageing, and consequently a remarkable decrease in resistivity in one phase and an increase in the other. This situation obviously contradicts the experimental observation, so we can conclude that no

LRO change is involved during ageing and that the small decrease in resistivity of both phases during ageing is due to a type of relaxation process that does not change LRO.

There is only one type of atomic rearrangement that does not change the degree of LRO: atomic rearrangement within the same sublattice of the ordered martensite, not between sublattices. This process is similar to the atomic diffusion process in ordered alloys, where effective (or net) atomic exchange occurs only within the same sublattice so as not to destroy the LRO²³. We show below that such an atomic rearrangement has a simple origin.

To simplify the discussion, we consider a two-dimensional structure. However, the conclusion is applicable to any three-dimensional martensite structure because only symmetry is relevant. Fig. 2a shows a two-dimensional A–B binary imperfectly ordered parent phase with four-fold symmetry. Because of the four-fold symmetry of the structure, the probability of finding a B atom about the A atom (or B atom) must possess the same four-fold symmetry (if not, four-fold crystal symmetry cannot be maintained), that is, $P_1^B = P_2^B = P_3^B = P_4^B$, $P_5^B = P_6^B = P_7^B = P_8^B$, and so on, where P_i^B ($i = 1, 2, 3, \dots$) are conditional probabilities, as explained in Fig. 2. The martensite has a lower symmetry, so the equilibrium occupation probability of atoms should also have the same symmetry as the structure about the A atom, that is, $P_1^B = P_3^B \neq P_2^B = P_4^B$ and $P_5^B = P_7^B \neq P_6^B = P_8^B$, as shown in Fig. 2c. The equalities are due to the centro-symmetry of the schematic martensite structure.

Now, let the parent phase shown in Fig. 2a transform into martensite. Because of the diffusionless nature of the transformation, all the probabilities must remain unchanged despite the symmetry change, as shown in Fig. 2b. That is, $P_1^B = P_2^B = P_3^B = P_4^B$ and $P_5^B = P_6^B = P_7^B = P_8^B$, and so on. But this high-symmetry configuration is no longer a stable configuration for the lower-symmetry martensite structure. Then during ageing, such a configuration gradually changes into the stable one, as shown in Fig. 2c. This process proceeds by atomic rearrangement or relaxation within the same sublattice. This is the only way that a martensite can lower its free energy without altering the average martensite structure (equilibrium phase). It is obvious that such an atomic rearrangement occurs most easily for the atoms close to the A atom (for example, atoms 1–8 shown in Fig. 2), and farther atoms make only a small contribution. This means that it is a type of short-range ordering (SRO) within the same sublattice. As shown above, this SRO originates from a requirement that the symmetry of the configuration of lattice imperfections in equilibrium conforms to

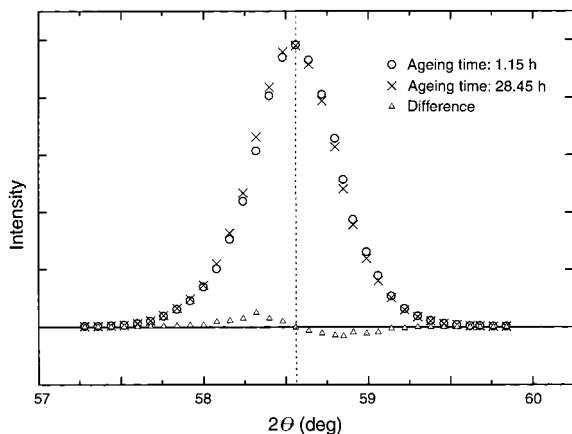


Figure 3 X-ray profiles of (242) Bragg reflection of single-crystal $\text{Au}_{52.5}\text{Cd}_{47.5}$ martensite after ageing for a short (1.15 h) and a long (28.45 h) time²¹. No perceptible change in peak position and integrated intensity was found during ageing. However, a delicate change in the symmetry of the profile was found to develop during ageing, as shown by the difference in the two profiles.

the crystal symmetry and is thus called symmetry-conforming short-range ordering. The stable SRO configuration for the parent is inherited by martensite during martensite formation, but this SRO becomes an 'unstable' configuration for the martensite because it differs from the martensite symmetry. Then atomic rearrangement occurs during ageing that results in a correct SRO symmetry for the martensite. This SRO is apparently different from previously proposed ones^{15–17} that lead to a change in LRO.

When the stabilized (or aged) martensite (Fig. 2c) is deformed, it changes into another domain (or twin) as a result of the accommodation of the strain. Because this twinning process is also diffusionless, the atomic occupation probabilities shown in Fig. 2c are inherited by the new domain, as shown in Fig. 2d, but this configuration is not stable for the new domain, which is shown in Fig. 2e, and a driving force that tries to restore the original domain (Fig. 2c) develops. When the external stress is released immediately after loading, this restoring force restores the new domain (Fig. 2d) to the original one (Fig. 2c) by de-twinning. By this mechanism the aged martensite can be deformed like a rubber. We propose that this is the origin of the RLB. If the stress is held for some time, the atomic configurations in Fig. 2d have enough time to change into a stable configuration (Fig. 2e), so no RLB occurs. This has also been found in previous experiments^{11,20}. It should be emphasized that this mechanism makes use of only two main features of the martensitic transformation—symmetry change and absence of diffusion—which are shared by all martensites. Thus it is a general mechanism applicable to any martensite.

When the stabilized martensite (Fig. 2c) is heated and transformed back (diffusionlessly) into the parent, the stable SRO configuration for the martensite is inherited by the parent (Fig. 2f). From the above-mentioned symmetry-conforming principle of SRO, it is obvious that Fig. 2f is not a stable configuration for the parent. From a thermodynamic point of view this corresponds to an increased reverse transformation temperature (known as A_s). This is the origin of the so-called martensite stabilization⁹, which has previously had no unified explanation.

The present model can be easily extended into disordered alloys by considering the existence of only one sublattice. In this case the present model reduces to Christian's model²⁴, which was later elaborated by Otsuka and Wayman¹. This model explained the rubber-like elasticity in disordered alloys such as In-Tl ^{25,26}.

We are able to verify the present model directly by comparing it with published X-ray diffraction measurements²¹ (Fig. 3). The invariance of the Bragg peak position (that is, lattice parameters)

and intensity indicates that atomic rearrangement during ageing does not change the average structure or LRO of martensite. The small change in shape of the Bragg peak with ageing reflects a delicate change in the environment of a given atom from Fig. 2b to Fig. 2c. Such a change leads to a small change in displacement field around a given atom, and results in a change in the shape of Bragg peaks. The change of the Bragg peak symmetry shows that the SRO configuration changes into the same symmetry as the martensite (the atomic structure of Au-Cd martensite is asymmetrical). The same result was also found in $\text{Au}_{50.5}\text{Cd}_{49.5}$ (ref. 20) and Au-Cu-Zn (ref. 22) alloys. So we can see that all of these previously unexplained experimental observations support the present model. The small decrease in resistivity in both martensite and the parent phase (Fig. 1) also provides support for the above mechanism, as it corresponds to the 'fine adjustment' or 'relaxation' within the same sublattice of the ordered structure after a sudden change in symmetry.

From Fig. 2 one can easily see that if there is perfect ordering, no lattice imperfections exist and thus there is no symmetry-conforming short-range ordering of imperfections. Thus, the existence of lattice imperfection is a necessary condition for the ageing effect and RLB. It follows that the more imperfections there are, the more pronounced is the effect. Experimentally⁵, it has been reported that the deviation from stoichiometry (increase in imperfections) in Au-Cd compounds markedly enhances RLB and the ageing effect, and quenching (increase in imperfections) always promotes RLB and the ageing effect.

Behind the above explanation of RLB and the ageing phenomenon is a principle of potentially more general significance: the symmetry-conforming principle of SRO. We believe that this principle is a key to understanding a wide spectrum of diffusion phenomena after a forced symmetry change either by a diffusionless transition or by an external field. RLB and the stabilization effect, as well as Zener ordering^{27,28} (due to a symmetry change by an external field), are just a few examples of this common origin. □

Received 19 May; accepted 29 July 1997.

- Otsuka, K. & Wayman, C. M. in *Review on the Deformation Behavior of Materials* Vol. 2 (ed. Feltham, P.) 81–172 (Freund, Tel Aviv, Israel, 1977).
- Christian, J. W. Deformation by moving interface. *Mater. Trans.* **13A**, 509–538 (1982).
- Cahn, R. W. Metallic rubber bounces back. *Nature* **374**, 120–121 (1995).
- Ölander, A. An electrochemical investigation of solid cadmium–gold alloys. *J. Am. Chem. Soc.* **56**, 3819–3833 (1932).
- Nakajima, Y., Aoki, S., Otsuka, K. & Ohba, T. The rubber-like behavior of ζ' (trigonal) martensite in $\text{Au-49.5 to 50.0 at\% Cd}$ alloys. *Mater. Lett.* **21**, 271–274 (1994).
- Miura, S., Maeda, S. & Nakanishi, N. Pseudoelasticity in Au-Cu-Zn thermoelastic martensite. *Phil. Mag.* **30**, 565–581 (1974).
- Rapacioli, R., Chandrasekaran, M. & Ahlers, M. in *Shape Memory Effects in Alloys* (ed. Perkins, J.) 365–378 (Plenum, New York, 1975).
- Sakamoto, H., Otsuka, K. & Shimizu, K. Rubber-like behavior in a Cu-Al-Ni alloy. *Scripta Metall.* **11**, 607–611 (1977).
- Tadaki, K., Otsuka, K. & Shimizu, K. Shape memory alloys. *Annu. Rev. Mater. Sci.* **18**, 25–45 (1988).
- Birnbaum, H. K. & Read, T. A. Stress induced twin boundary motion in $\text{AuCd } \beta'$ and β'' alloys. *Trans. AIME* **218**, 662–669 (1960).
- Lieberman, D. S., Schmetling, M. A. & Karz, R. S. in *Shape Memory Effects in Alloys* (ed. Perkins, J.) 203–244 (Plenum, New York, 1975).
- Zangwill, A. & Bruinsma, R. Origin of martensitic pseudoelasticity. *Phys. Rev. Lett.* **53**, 1073–1076 (1984).
- Abu Arab, A. & Ahlers, M. The stabilization of martensite in Cu-Zn-Al alloys. *Acta Metall.* **36**, 2627–2638 (1988).
- Tadaki, K., Okazaki, H., Nakata, Y. & Shimizu, K. Atomic configuration studies by ALCHEMI and X-ray diffraction of a stabilized M18R martensite in a β phase Cu-Au-Zn alloy. *Mater. Trans. Jpn Inst. of Metals* **31**, 941–947 (1990).
- Ahlers, M., Barcelo, G. & Rapacioli, R. A model for the rubber-like behavior in Cu-Zn-Al martensites. *Scripta Metall.* **12**, 1075–1078 (1978).
- Marukawa, K. & Tsuchiya, K. Short-range ordering as the cause of the rubber-like behavior in alloy martensites. *Scripta Metall.* **32**, 77–82 (1995).
- Suzuki, T., Tonokawa, T. & Ohba, T. Role of short-range order in martensitic transformation. *J. Phys. III* **5**(C8), 1065–1070 (1995).
- Cahn, J. W. Thermodynamic and structural changes in deformation twinning of alloys. *Acta Metall.* **25**, 1021–1026 (1977).
- Barcelo, G., Rapacioli, R. & Ahlers, M. The rubber effect in Cu-Zn-Al martensite. *Scripta Metall.* **12**, 1069–1074 (1978).
- Murakami, Y., Nakajima, Y., Otsuka, K. & Ohba, T. Effect of existing twin boundaries in martensite on the rubber-like behavior and ageing effect in Au-Cd alloys. *J. Phys. III* **5**(C8), 1071–1076 (1995).
- Ohba, T., Otsuka, K. & Sasaki, S. Study of rubber-like behavior in a $\text{Au-47.5 at\% Cd}$ alloy by synchrotron-orbital radiation. *Mater. Sci. Forum* **56–58**, 317–322 (1990).
- Ohba, T., Finlayson, T. & Otsuka, K. Diffraction profile change in Au-Cu-Zn alloy with aging. *J. Phys. III* **5**(C8) 1083–1086 (1995).
- Hagel, W. C. in *Intermetallic Compounds* (ed. Westbrook, J. H.) 377–404 (Wiley, New York, 1967).
- Christian, J. W. *The Theory of Transformations in Metals and Alloys* (Pergamon, Oxford, 1965).

25. Burkart, M. W. & Read, T. A. Diffusionless phase change in the indium–thallium system. *Trans. AIME* **197**, 1516–1523 (1953).
26. Basinski, Z. S. & Christian, J. W. Crystallography of deformation by twin boundary movement in indium–thallium alloys. *Acta Metall.* **2**, 101–116 (1954).
27. Zener, C. Stress induced preferential orientation of pairs of solute atoms in metallic solid solution. *Phys. Rev.* **71**, 34–38 (1947).
28. Birkenbeil, H. J. & Cahn, R. W. Induced magnetic anisotropy created by magnetic and stress annealing of iron–aluminium alloys. *Proc. Phys. Soc.* **79**, 831–847 (1962).
29. Miyazaki, S. & Otsuka, K. in *Shape Memory Alloys* (ed. Funakubo, H.) 116–175 (Gordon and Breach, New York, 1987).

Acknowledgements. We thank T. Suzuki, T. Ohba, H. Horiuchi, M. Kogachi and Y. Murakami for helpful discussions and critical reading of the manuscript; T. Ishii for technical assistance; and T. Ohba for allowing us to use his published diffraction data. The work was supported by a Grant-in-Aid for Scientific Research on Priority Area of Phase Transformations (1997–1999) from the Ministry of Education, Science and Culture of Japan. X.R. is a recipient of a JSPS Fellowship.

Correspondence and requests for materials should be addressed to K.O. (e-mail: otsuka@mat.ims.tsukuba.ac.jp).

Bending and buckling of carbon nanotubes under large strain

M. R. Falvo*, G. J. Clary†, R. M. Taylor II†, V. Chi†, F. P. Brooks Jr†, S. Washburn* & R. Superfine*

* Department of Physics and Astronomy, †Department of Computer Science, University of North Carolina, Chapel Hill, North Carolina 27599, USA

The curling of a graphitic sheet to form carbon nanotubes¹ produces a class of materials that seem to have extraordinary electrical and mechanical properties². In particular, the high elastic modulus of the graphite sheets means that the nanotubes might be stiffer and stronger than any other known material^{3–5}, with beneficial consequences for their application in composite bulk materials and as individual elements of nanometre-scale devices and sensors⁶. The mechanical properties are predicted to be sensitive to details of their structure and to the presence of defects⁷, which means that measurements on individual nanotubes are essential to establish these properties. Here we show that multiwalled carbon nanotubes can be bent repeatedly through large angles using the tip of an atomic force microscope, without undergoing catastrophic failure. We observe a range of responses to this high-strain deformation, which together suggest that nanotubes are remarkably flexible and resilient.

Two mechanical properties of a material are of principal interest here: the small-strain elastic modulus and the material strength. The basal-plane elastic modulus of graphite, the largest of any known material, confers on the nanotube its predicted extraordinary stiffness^{3,4,24}. These expectations are consistent with reports of transmission electron microscope (TEM) measurements taken in the small-strain limit⁵, and atomic force microscope (AFM) measurements which were extended beyond the linear elastic regime⁸. Although the high stiffness is predicted to provide an improvement over existing materials, it is the nanotube's unusual strength which is its most distinguishing property. The unusual strength arises from the high stiffness, expected to be consistent with the modulus of graphite, combined with extraordinary flexibility and resistance to fracture. It is this latter feature which will distinguish nanotubes from graphitic fibres as an engineering material. Two theoretical approaches to understanding and predicting the large-strain behaviour of carbon nanotubes are atomistic molecular-dynamics simulations and continuum mechanics. The remarkable correspondence between the two approaches and TEM observations has been appreciated recently in the case of the smallest single- and double-wall tubes under flexural strains which produce buckling deformations⁹. Large multiwalled nanotubes (MWNTs) present challenges to both theoretical approaches in that molecular-dynamics simulations are limited in the size of the system that can be modelled, and continuum mechanics has been most successful in the thin-shell limit.

Knowledge of the large-strain behaviour of carbon nanotubes relies largely on the TEM observations of as-deposited, small-diameter tubes. Typically, many small kinks are seen on the compressed side of an otherwise smooth bend in a MWNT^{10,11}; molecular-dynamics simulations predict this kinking to be reversible even under very severe bending¹². This is consistent with indirect evidence of reversible buckling of a carbon nanotube while attached to an AFM probe⁸. This indicates a fibre of unusual strength. Until now, experimental evidence of the structure of bent nanotubes has relied largely on finding nanotubes which have been bent either during the deposition procedure or through the distortion of the substrate^{13,14}. Ideally, experiments should be performed where the full history of the strain, including the initial condition of the tube and the reversibility of the structure, is known. Using a unique interface for atomic force microscopes, we have performed intricate bending of MWNTs to large strains.

Samples were prepared by solvent evaporation on mica substrates of an ethanol solution of raw carbon soot (produced from the carbon-arc technique¹⁵). The carbon nanotubes were imaged and manipulated under ambient conditions. The 'Nanomanipulator' AFM system, designed for manipulation, comprises an advanced visual interface, teleoperation capabilities for manual control of the AFM tip and tactile presentation of the AFM data (Discoverer, Topometrix Inc., Santa Clara, CA)^{16,17}. Initially, the tubes are straight with no apparent structural defects. The AFM tip is used to apply lateral stresses at locations along the tube to produce translations and bends. The friction between the tube and the substrate pins the tube in its strained configuration for imaging. On some samples, however, the tube/substrate friction is too low to maintain the highly strained nanotube, and relaxation of the tube is observed during AFM imaging.

A carbon nanotube, pinned at one end by carbon debris, was bent into many configurations. The inset of Fig. 1a shows the tube in its original adsorbed position and orientation. The abrupt vertical step on the left side of the tube was used as a reference mark ($s = 0$, where s is the measured distance along the tube centre line) for feature locations. The nanotube was taken through a series of 20 distinct manipulations, alternately bending and straightening the tube at various points. We present images from this sequence to highlight specific features (Fig. 1a–d). Along with AFM topographic data, we also show the tube height along the tube's centre as determined by the cores method^{18,19}, along with the calculated curvature (Fig. 1e–h). We will refer to raised points on the tube as 'buckles', consistent with the increase in height expected from the collapse of a shell in response to bending. We observe two behaviours in this sequence: small, regularly distributed buckles at regions of small curvature and large deformation at high curvatures.

We focus first on the regularly spaced buckles occurring in the more gradually bent region (Fig. 1a, lower right). Figure 1b and c shows the tube as it is bent in opposite directions. The location of the buckles has shifted dramatically, with the buckles of Fig. 1c appearing in regions which had been featureless, and the buckles of Fig. 1b largely disappearing. The buckles appear with a characteristic interval independent of their absolute position along the tube. This suggests the buckling is reversible, intrinsic to the nanotube and not mediated by defects. The strong correlation between tube curvature and the location of buckles confirms their role in reducing curvature-induced strain. Buckling in bent shells are well known in continuum mechanics²⁰ where two modes of deformation result from pure bending stresses. The Brazier effect causes the circular cross-section of the tube to become more 'ovalized' uniformly over the whole tube length²¹ as the bending curvature increases. Bifurcation, on the other hand, leads to a periodic, low-amplitude rippling of the tube wall on the inside of the bend (the portion under compression). Experiment and theory show²¹ that both modes often occur simultaneously, and lead to single-kink collapse as the bending moment reaches its maximum critical value. The axial

bifurcation (rippling) wavelength is given by²²:

$$\lambda = \frac{2\pi}{(12(1-\nu^2))^{1/4}} (r_0 t)^{1/2} \approx 3.5(r_0 t)^{1/2}$$

where r_0 is the tube radius and t is the thickness of the tube wall; we have set the Poisson ratio $\nu = 0.3$ as expected for carbon nanotubes. The tube shown in Fig. 1 has a radius of 10.5 nm. If we assume that the buckles are collapsed regions with the height determined by the complete flattening²³ of the tube as constrained by the inner radius, we obtain $t = 2.6$ nm and $\lambda = 18.3$ nm. We have plotted a histogram of buckle intervals (Fig. 1d inset) and the average of the Fourier transforms (Fig. 1c, inset) for the various bent configurations. The mean value of the interval distribution (68 nm) coincides with the

peak in the Fourier transform at 0.015 nm^{-1} . These values are roughly 4 times the predicted value of λ . We note two points. First, λ is for a fundamental buckling mode and modes of higher order are predicted²². Second, the buckling models have been developed for thin wall sections, $r_0/t \gg 10$, but in our sample $r_0/t \approx 4.0$.

Under the largest bending strains this tube buckles more severely. This is seen in the large height anomalies observed at the top right of Fig. 1a and b. The height of the nanotube to the right of the step is ~ 21 nm. As shown in Fig. 1f, the height of the highest buckle is almost twice the diameter of the tube. In the simplest model of a thin-walled tube collapsed into a ribbon, the width of the ribbon is πr_0 . In the absence of elastic stretching of the tube wall transverse to the tube axis, this is the widest a collapsed tube can be. We have

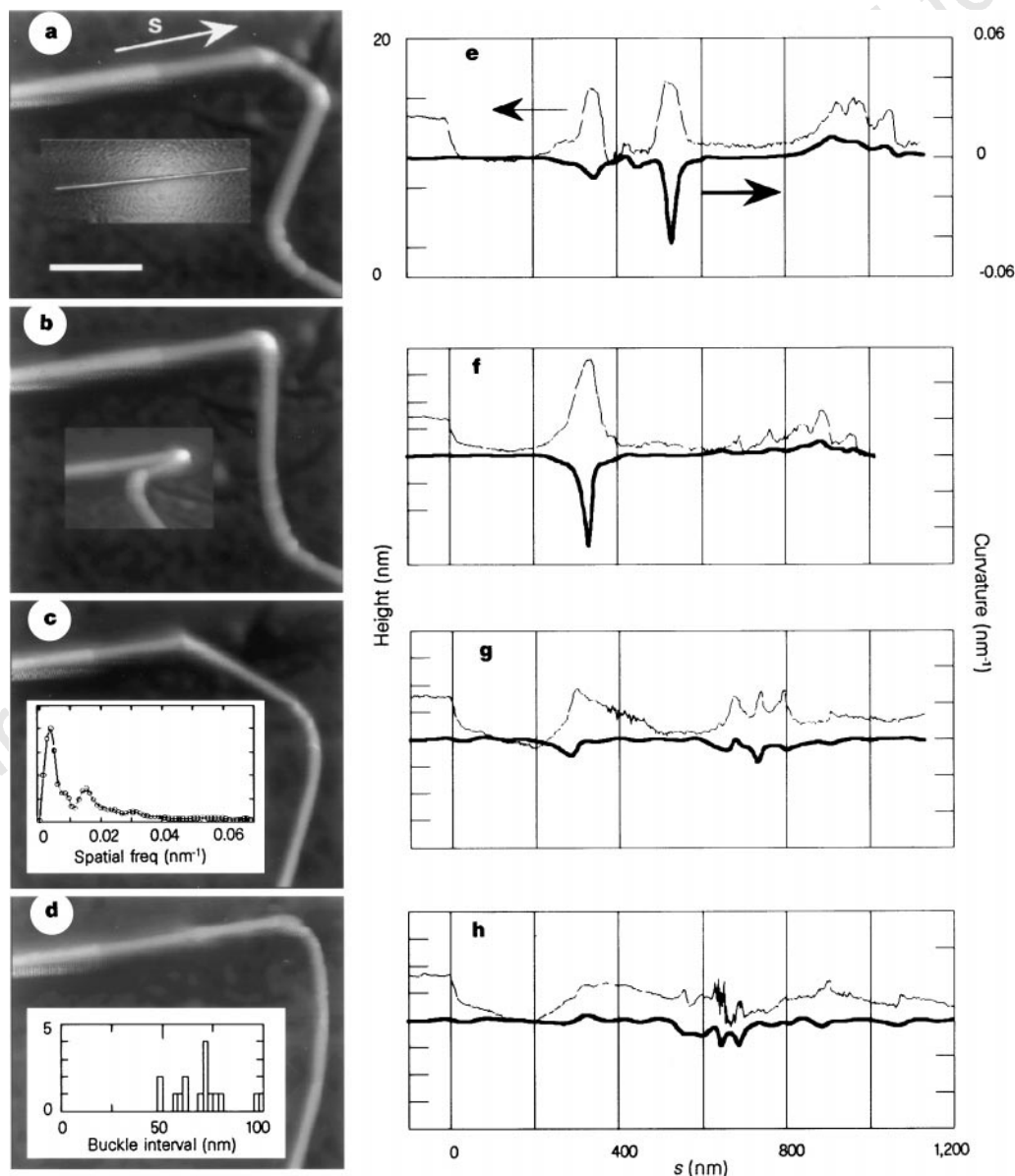


Figure 1 Curvature and height of buckles along a bent carbon nanotube. The white scale bar (in **a**) represents 300 nm and all figures are to the same scale. A 20-nm-diameter tube was manipulated (with an AFM) from its straight shape (**a** inset) into several bent configurations (**a**–**d**). The height and curvature of the bent tubes along its centreline (indicated by the arrow in **a**) is shown in **e**–**h**. The upper trace in each graph depicts the height relative to the substrate; the lower trace depicts the curvature data. Height values are relative to substrate height. The 'ripple'-like buckles occur between $s = 600$ and $1,200$ nm and correspond to the tube in the lower right portion of the images in each case. We note that these buckles migrate as the tube is manipulated into different configurations. Twice during the manipu-

lation sequence the tube is bent to $\sim 180^\circ$ (**b** inset). The appearance and disappearance of the ripple buckles, as well as the severe buckle at $s \approx 500$ nm (**e**–**f**), suggest elastic reversibility. The large buckle at $s \approx 325$ nm (**e**, **f**) retains its raised topographical features even after straightening (**g**, **h**), suggesting that damage has occurred at this point; but the tube does not fracture. The distance between 'ripple' buckles was determined for the various bent configurations (**a**–**d**). The average of the buckle interval histogram (**d**, inset) and the average of the Fourier transforms (**c** inset) for a wide range of bent configurations establish the dominant interval as 68 nm. The low-frequency peak in the Fourier transform correspond to the long-wavelength tube distortion over the entire bend region.

observed buckle heights of up to $4r_0$. This suggests that either the tube is rising up off the sample surface at these buckling points, or that the tube is rupturing and we are imaging torn sections projecting upwards. As the tube is straightened, these sections do not return to their initial height, and the unusual projections return on rebending. The radii of curvature obtained in these bends, ~ 20 nm (Fig. 1b inset), is the same as the diameter of the tube. It is remarkable that even though the tube sustains damage after twice bending and straightening the section, it does not separate.

Although the nanotube shown in Fig. 1 appeared to suffer permanent distortion, we have observed other cases where the nanotube survives large strains without gross damage. In Fig. 2a, a ~ 800 -nm-long tube with $r_0 \approx 6.8$ nm was bent repeatedly to large strain with no permanent distortion of the tube topography. The left side of the tube was translated towards the top of the figure, bending the tube back on itself (Fig. 2b). The analysis of the ridge and the curvature (Fig. 2e–g) reveals behaviour distinct from the previous case. Figure 2b shows a bend with a radius of curvature (Fig. 2e) comparable to that of the large bend in Fig. 1c, but with a change in height of < 3 nm. Figure 2c shows the tube with $r_c = 20$ nm $\approx 3r_0$, with a change in height of 4 nm (Fig. 2f). In Fig. 2d, the tube is bent back onto itself in the other direction. This nanotube behaves very differently from the nanotube of Fig. 1 because it bends with only a single broad flattening, with no indication of periodic buckling at any curvature. The difference between the two tubes may lie in the

r_0/t ratio. If we estimate the wall thickness from the tube flattening, we obtain $t = 3.25$ nm and $r_0/t = 2.1$ instead of $r_0/t = 4.0$ for the tube of Fig. 1. The height of the tube of Fig. 2 always remains within the bounds of a flattened tube, even after being bent through the sequence of curvature radii: $+20$ nm, $+40$ nm, where positive curvature has been assigned to Fig. 2b. The maximal local strain in this tube can be estimated from $\epsilon = r_0/r_c \approx 16\%$ on the outside tube surface, ($r_c =$ radius of curvature of bend), with a corresponding compressive strain on the inner surface. The apparent lack of catastrophic damage of the tube under such large strains is remarkable. We can speculate on two ways in which the strain has been accommodated. First, it has been proposed⁷ that MWNTs contain significant concentrations of defects. It might be that defects in this tube are arranged in an incoherent fashion such that they separate under tensile stress, and slide over each other reversibly under compressive stress. Second, strain might be accommodated within a severely distorted, but otherwise connected, graphene sheet. Molecular-dynamics simulations of single-walled tubes under tensile strain have shown, under certain conditions, breaking strains as large as 30% (ref. 24).

Thus carbon nanotubes are extraordinarily flexible under large strains, and resist failure under repeated bending. We have observed reversible, periodic buckling of nanotubes consistent with calculations extrapolated from continuum mechanics. Local strains as large as 16% can be sustained without separating a nanotube. In fact, we have never observed the separation of a tube by the repeated application of bending stresses. The amount of strain that we have been able to apply to a carbon nanotube has been limited by the surface forces which hold the tube in place. These results show that the carbon nanotube is a material with extraordinary strength. $\square$

Received 6 June; accepted 5 August 1997.

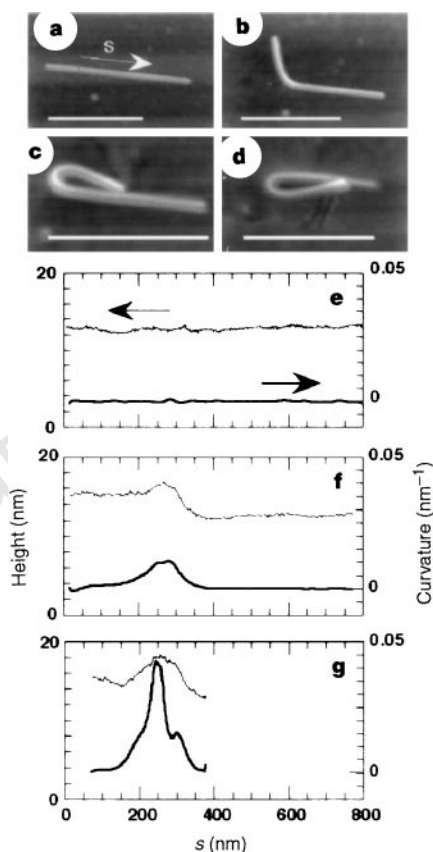


Figure 2 Carbon nanotube in highly strained configuration. The white scale bars at the bottom of each panel a–d represent 500 nm. **a**, The original adsorbed shape of the tube, 10.5 nm in diameter and 850 nm long. The tube is bent in steps, first upwards (**b**), until it bends all the way back onto itself (**c**). The curvature is 0.045 nm^{-1} (radius of curvature ~ 20 nm) which indicates the strain along the inside and outside of the tube bend is $\sim 16\%$. **d**, The tube is then bent all the way back the other way onto itself, to a final curvature similar to that in **c**. Panels **e–g** correspond to **a–c** respectively, and show the height and curvature along the tube in its different configurations. This tube did show signs of irreversible damage when straightened between bends, but never 'kinked' or fractured under bending stress.

- Iijima, S. Helical microtubules of graphitic carbon. *Nature* **354**, 56–58 (1991).
- Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
- Robertson, D. H., Brenner, D. W. & Mintmire, J. W. Energetics of nanoscale graphitic tubules. *Phys. Rev. B* **45**, 12592–12595 (1992).
- Sawada, S. & Hamada, N. Energetics of carbon nano-tubes. *Solid State Commun.* **83**, 917–919 (1992).
- Treacy, M. M. J., Ebbesen, T. W. & Gibson, J. M. Exceptionally high Young's modulus observed for individual carbon nanotubes. *Nature* **381**, 678–680 (1996).
- Dai, H., Hafner, J. H., Rinzler, A. G., Colbert, D. T. & Smalley, R. E. Nanotubes as nanoprobe in scanning probe microscopy. *Nature* **384**, 147–150 (1996).
- Zhou, O. *et al.* Defects in carbon nanostructures. *Science* **263**, 1744–1747 (1994).
- Lieber, C. Nanotube derived materials. *Bull. Am. Phys. Soc.* **42**, 591 (1997).
- Yakobson, B. I., Brabec, C. J. & Bernholc, J. Nanomechanics of carbon tubes: instabilities beyond the linear response. *Phys. Rev. Lett.* **76**, 2511–2514 (1996).
- Ruoff, R. S. & Lorents, D. C. Mechanical and thermal properties of carbon nanotubes. *Carbon* **33**, 925–929 (1995).
- Despres, J. E., Daguerre, D. & Lafdi, K. Flexibility of graphene layers in carbon nanotubes. *Carbon* **33**, 87–92 (1995).
- Iijima, S., Brabec, A., Maiti, A. & Bernholc, J. Structural flexibility of carbon nanotubes. *J. Chem. Phys.* **104**, 2089–2092 (1996).
- Ebbesen, T. W., in *Carbon Nanotubes: Preparation and Properties* (ed. Ebbesen, T. W.) 225–248 (CRC, Boca Raton, 1997).
- Ruoff, R. S. & Kadish, K. M. *Fullerenes; Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials Vol. 2* (Electrochem. Soc., Pennington, NJ, 1995).
- Ebbesen, T. W. & Ajayan, P. M. Large scale synthesis of carbon nanotubes. *Nature* **358**, 220–222 (1992).
- Taylor, R. M. II *et al.* *The Nanomanipulator: A Virtual Reality Interface for a Scanning Tunneling Microscope*. (ed. J. T. Kajiya) 127–134 (Association for Computing Machinery, New York, 1993).
- Finch, M. *et al.* *Surface Modification Tools in a Virtual Environment Interface to a Scanning Probe Microscope*. (ed. P. Hanrahan & J. Winget) 13–18 (Association for Computing Machinery, New York, 1995).
- Pizer, S. M., Eberly, D., Morse, B. S. & Fritsch, D. S. Zoom-invariant vision of figural shape: the mathematics of cores. *Comput. Vision Image Understand.* (in the press).
- Fritsch, D. S., Eberly, D., Pizer, S. M. & McAuliffe, M. J. Simulated cores and their applications in medical imaging. *Information Processing in Medical Imaging, Proc. IPMI '95* (ed. Y. Bazals, C. Barillot and R. Di Paola) 365–368 (Kluwer, Dordrecht, The Netherlands (1994)).
- Landau, L. D. & Lifshitz, E. M. *Theory of Elasticity* (Pergamon, Oxford, 1986).
- Ju, G. T. & Kyriakides, S. Bifurcation and localization instabilities in cylindrical shells under bending – II. Predictions. *Int. J. Solids Structures* **29**, 1143–1171 (1992).
- Axelrad, E. L. On local buckling of thin shells. *Int. J. Non-Linear Mech.* **20**, 249–259 (1985).
- Chopra, N. G. *et al.* Fully collapsed carbon nanotubes. *Nature* **377**, 135–138 (1995).
- Yakobson, B. I., Campbell, M. P., Brabec, C. J. & Bernholc, J. High strain rate fracture, and C-chain unraveling in carbon nanotubes. *Comput. Mater. Sci.* **8**, 341–348 (1997).
- Lu, J. P. Elastic properties of carbon nanotubes and nanoropes. *Phys. Rev. Lett.* **79**, 1297–1300 (1997).

Acknowledgements. We thank O. Zhou for providing the carbon nanotube soot and S. Paulson for help with sample preparation. This work was supported by the National Science Foundation, National Institutes of Health National Centers for Research Resources, Topometrix, Inc. and Silicon Graphics, Inc.

Correspondence and requests for materials should be addressed to R.S. (e-mail: rsuper@physics.unc.edu).

Synthesis of cadmium sulphide superlattices using self-assembled bacterial S-layers

Wayne Shenton^{*†}, Dietmar Pum^{‡†}, Uwe B. Sleytr[‡] & Stephen Mann^{*}

^{*} School of Chemistry, University of Bath, Bath BA2 7AY, UK

[‡] Center for Ultrastructure Research and Ludwig Boltzmann Institute for Molecular Nanotechnology, Universität für Bodenkultur Wien, Vienna, Austria

[†] These authors contributed equally to this work

Methods for organizing materials at the nanometre scale have advanced tremendously in recent years^{1,2}. One important objective is the synthesis of patterned arrays of inorganic nanocrystals^{3–6}, whose optical, electronic and magnetic properties might find technological uses, for example as memory elements. Techniques such as colloidal crystallization^{7–9}, monolayer deposition^{10–12}, multilayer casting¹³, molecular crosslinking^{14,15}, the use of complementary interactions^{16,17} and the synthesis of nanoparticles in patterned etch pits¹⁸ have been used to organize nanocrystals into superlattices. Here we describe the use of bacterial S-layers—self-assembled, two-dimensionally ordered films of proteins that feature in many bacterial cell walls—as templates for the *in situ* nucleation of ordered two-dimensional arrays of cadmium sulphide nanocrystals about 5 nm in size. Nucleation of the inorganic phase is confined to the pores between subunits in the S-layers. Two-tier stacks of nanoparticles can be formed in the presence of double-layered protein crystals. The structural diversity of S-layers^{19,20}, their ease of self-assembly on a wide range of substrates and the potential for surface chemical modification suggest that this approach could be exploited to offer a wide range of ordered nanoparticle arrays.

With a few exceptions, S-layers represent an almost universal feature of archaeobacterial cell envelopes, and have been identified in many different species of nearly every taxonomic group of walled eubacteria^{19,20}. S-layers are two-dimensional crystalline structure of single protein or glycoprotein monomers (relative molecular mass, M_r , 40,000 to 200,000), and exhibit either oblique ($p1$, $p2$), square ($p4$) or hexagonal ($p3$, $p6$) lattice symmetry with spacings between the morphological units in the range 3–30 nm (ref. 20). Most S-layers are 5–15 nm in thickness, possess pores of identical size and morphology in the 2–6 nm range, and have inner and outer surfaces that are markedly different in topography and physicochemical properties²¹. Although considered primarily as a filtration and structural matrix, S-layers have been implicated in cell wall biomineralization^{22,23}, suggesting a possible biomimetic role for these proteins as templates in materials chemistry.

Our approach to using S-layer templates for inorganic superlattice construction is illustrated in Fig. 1 (see Methods for details). Figure 2a shows a typical monolayer of the self-assembled S-layer of *Bacillus stearothermophilus* NRS2004/3a variant 1. The sample was negatively stained with uranyl acetate, which penetrates the nanoporous protein structure to reveal an oblique lattice (space group, $p1$; $a = 9.8$ nm, $b = 7.5$ nm, $\theta = 80^\circ$). The protein fine structure, which appears white in the stain exclusion pattern, is imaged to a resolution of ~ 2 nm. Similar studies with recrystallized S-layers of *B. sphaericus* CCM 2177 revealed a square lattice with a unit cell length of 13 nm (data not shown). Mineralization of the exposed inner face of the self-assembled proteins (see Methods) produced S-layers that were decorated with an organized array of CdS nanoparticles in register with the underlying periodicity of the protein

lattice (Fig. 1a). Thus, an oblique inorganic superlattice ($a = 9.8$ nm, $b = 7.5$ nm, $\theta = 80^\circ$) was formed using self-assembled *B. stearothermophilus* NRS 2004/3a variant 1 S-layers (Fig. 2b), whereas a square superlattice ($a = 13$ nm) was templated on the inner face of recrystallized *B. sphaericus* CCM 2177 S-layers (Fig. 2c). In both systems, the CdS clusters were discrete, relatively uniform in size (mean values; 5 nm (oblique lattice); 8 nm (square lattice)) and rounded in shape. Mineralized arrays with domain sizes up to 1 μ m were observed.

Electron-diffraction patterns and high-resolution lattice images obtained with transmission electron microscopy (TEM) indicated that the CdS nanoparticles were crystalline with d spacings corresponding to the zinc blende structure (d spacings, 0.336, 0.206, 0.176, 0.133, 0.118 nm). The patterns showed no preferential orientation, indicating that the particles were not crystallographically aligned on the hydrophilic inner face of the assembled S-layer subunits. Although the presence of the inorganic array dominated the TEM images, a faint pattern of the protein structure with its typical handedness of connecting arms was visible in the background. This observation suggests that Cd(II) also stains the S-layer template, and that the CdS nanoparticles are deposited within the network of nanopores which are located between the protein subunits. This was confirmed by the digital diffraction patterns which showed only low-order reflections for the reciprocal oblique lattice of the S-layer monolayer.

CdS mineralization of the exposed outer face of *B. stearothermophilus* NRS2004/3 variant 1 S-layers (Fig. 1b) produced arrays of discrete 2–3 nm sized CdS particles that were preferentially deposited on the S-layer template (Fig. 2d). Electron diffraction gave broadened ring patterns with d spacings consistent with the zinc blende structure. Close examination of the micrographs revealed some periodicity in the CdS array but the extent of replication was significantly less than that observed for the nanoparticles formed on the inner surface of the protein template (Fig. 2b). In both cases, the results suggest that the inorganic clusters are located within the nanopores of the protein matrix. The inter-subunit spaces are larger on the inner face because of the higher surface corrugation compared with the relatively flat outer surface²⁴. This difference in topography could account for the larger size of the CdS nanoparticles, as well as the increased fidelity in the periodicity of the inorganic arrays deposited on S-layers recrystallized with the inner face exposed. In addition, the inner face has a net negative charge and is less hydrophobic than the charge-neutral outer surface, suggesting that electrostatic binding of Cd(II) ions might also be important in site-directed nucleation.

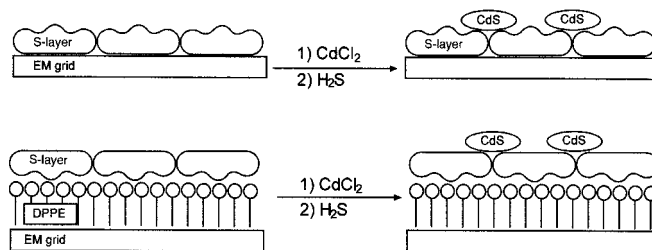


Figure 1 Substrate-assisted self-assembly of bacterial S-layer monolayers used for biocrystal templating of CdS superlattices. **a**, Recrystallization of S-layers on electron microscope grids produces a nanoporous monolayer with the corrugated negatively charged inner face exposed to the external medium. **b**, Corresponding orientation of a recrystallized S-layer formed under a Langmuir monolayer of the lipid, dipalmitoylphosphatidylethanol (DPPE) and transferred by horizontal dipping to an electron microscope grid. By this method, the charge-neutral outer face of the self-assembled S-layer is exposed²⁰. In both systems, addition of a Cd(II) solution, followed by reaction with H₂S, results in site-specific nucleation in the inter-subunit regions to produce an organized array of CdS nanoparticles.

We also investigated the synthesis of CdS nanocrystal arrays using dispersed S-layer structures formed by recrystallization from bulk aqueous solution (see Methods). A well-defined square superlattice ($a = 13$ nm) of 5-nm-sized CdS crystallites was deposited on S-layers of *B. sphaericus* CCM2177 after 100 min of exposure to H_2S in the presence of Cd(II) (data not shown). In contrast, mineralization of the dispersed microstructures formed in suspensions of *B. stearotheophilus* NRS2004/3a variant 1 S-layers produced characteristic stripe patterns of organized CdS nanocrystals (Fig. 3A). Energy-dispersive X-ray analysis showed the presence of Cd and S throughout the S-layer sheets, and electron diffraction and high-resolution lattice imaging indicated that the individual CdS nanoparticles within the striped regions were non-oriented, ~ 5 -nm-sized single crystals with the zinc blende structure. The most common stripe pattern consisted of fringes, 16 nm in width and spaced at ~ 32 nm, which were aligned parallel to the longer unit-cell vector of the underlying S-layer lattice (Fig. 3A). Digital diffraction patterns obtained from these images indicated that the stripe patterns were Moiré fringes which originated from the superimposition of two oblique CdS/protein superlattices with the longer base vector in common (Fig. 3B). The diffraction patterns showed mirror symmetry, indicating that the constituent S-layer/CdS monolayers were in a back-to-back orientation within a double-layered microstructure. In addition, the additional low-order reflection at 0.25 times the spacing of the longer reciprocal base vector (Fig. 3B) originates from the difference vector between the two shorter reciprocal base vectors, and corresponds to a repeat distance of 30 nm (4×7.5 nm) in the Moiré pattern. In this particular arrangement, the two associated back-to-back S-layers are in accurate register at every fourth row of the double-layered microstructure (Fig. 3C).

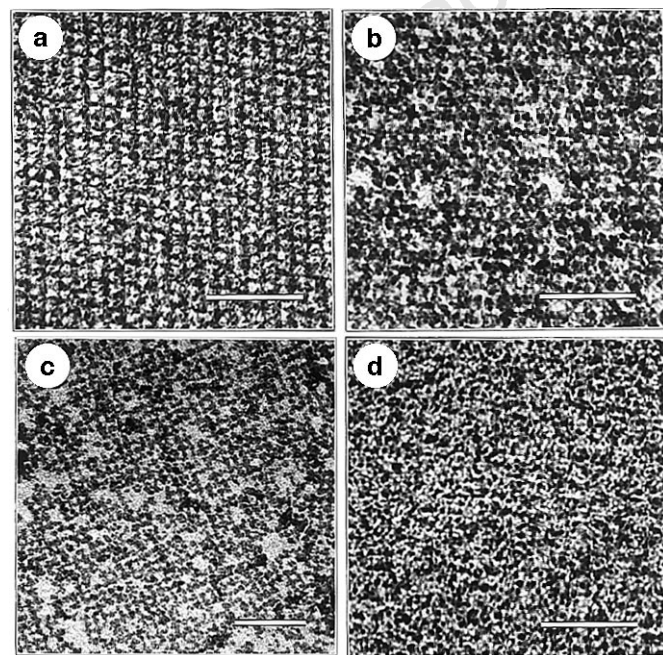


Figure 2 **a**, TEM micrograph of a negatively stained self-assembled S-layer of *B. stearotheophilus* NRS 2004/3a variant 1; scale bar, 60 nm. **b**, As for **a** but without staining and after CdS mineralization of the exposed inner face. The TEM image shows an oblique periodic array of uniform 5-nm-size CdS nanocrystals; scale bar, 60 nm. **c**, TEM micrograph showing a square CdS superlattice on the inner surface of a self-assembled *B. sphaericus* CCM2177 S-layer; scale bar, 100 nm. **d**, TEM micrograph of the mineralized outer face of an S-layer from *B. stearotheophilus* NRS2004/3a variant 1, showing CdS nanoparticles associated with the protein surface. The array shows weak periodicity (more readily observed by viewing the micrograph on its side); scale bar, 60 nm.

The data indicate that the double-layered protein template can be used to produce a stacked inorganic–organic assembly, consisting of an ordered two-tier arrangement of CdS nanocrystalline arrays. Previous experiments with polycationic ferritin, which is a topographical marker for net negatively charged domains on S-layer lattices, showed that the inner faces of the two constituent layers of the double-layered sheets were juxtaposed²⁵. Thus, inorganic precipitation directly onto the exposed surfaces of the dispersed sheets should replicate the nanoporous lattice of the flat outer faces. As this face was relatively ineffective as a template when mounted as a lipid/protein monolayer on TEM grids (Fig. 2d), mineralization of the double-layered structure possibly involves Cd(II) penetration of the architecture followed by *in situ* sulphidation within the confined spaces of the corrugated internalized inner faces. Further work is required to determine the precise location of the CdS nanocrystallites in these hierarchical materials.

Although the inorganic nanoparticles are not crystallographically orientated within the S-layer matrix, the preferential deposition of CdS in association with the dispersed S-layers, rather than in bulk solution, suggests that molecular interactions between the functional groups of the protein surface and Cd(II) ions in solution could be important for the construction of the superlattice arrays. As previous studies have shown that S-layer proteins can be chemically modified with many different types of functional groups and molecules²⁶, tailoring the surface properties of these biocrystal templates to wet chemical synthesis could have significant advantages over metallization procedures²⁷ in fabricating a wide range of organized inorganic nanomaterials with semiconducting, magnetic or electronic properties. In particular, there is at present great interest in the formation of metal cluster arrays for coulomb charging behaviour in nanoelectronic digital circuits²⁸. The strin-

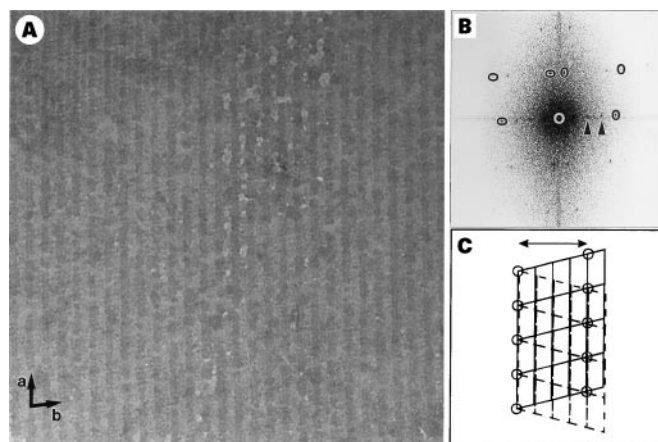


Figure 3 **A**, TEM micrograph of a mineralized double-layered sheet of *B. stearotheophilus* NRS2004/3a variant 1 S-layers, showing organized array of CdS nanoparticles and 32-nm stripe pattern due to Moiré effects; scale bar, 100 nm. **B**, Corresponding digital diffraction pattern showing superimposed oblique lattices with the longer reciprocal vector in common. The two reciprocal lattices are distinguished by either 0 or O symbols placed around their constituent reflections. The white circle corresponds to the centre spot and is the origin for both lattices. The composite pattern has mirror symmetry and additional lower-order reflections (arrow heads) spaced at 0.25 times the length of the longer reciprocal base vector. **C**, Sketch showing the relationship between two S-layers in back-to-back orientation within the double-layered sheet. The oblique real space lattices are coaligned on the longer axis and in register every four rows along the shorter base vector (arrow).

gent requirement for particle size monodispersity in such applications might be circumvented by directly synthesizing the nanoparticles in the presence of a suitable patterning agent. For this reason, the chemical synthesis of gold superlattices using S-layer templates is the principal objective of current and future work. □

Methods

S-layer proteins were extracted from isolated bacterial cell walls of *B. stearothermophilus* NRS 2004/3a variant 1 and *B. sphaericus* CCM 2177 as described previously²⁵. S-layer subunits were reassembled on carbon-coated formvar-covered nickel TEM grids by placing the grids onto the surface of a 1 mg ml⁻¹ buffered solution of the protein monomers for 2 h, after which the grids were removed by horizontal lifting, washed with distilled water, and the recrystallized S-layers chemically fixed with glutaraldehyde²⁹. Recrystallized S-layers were mineralized by placing TEM grids containing unstained protein monolayers onto drops of 10 mM CdCl₂ solution for 2 h. The grids were then lifted and surplus solution removed by blotting before being placed in a vacuum dessicator containing a dish of acidified Na₂S solution. Slow reaction of the hydrated samples with H₂S occurred for up to 2 d, after which the grids were removed and directly transferred into the electron microscope. Dispersions of self-assembled S-layers from either *B. stearothermophilus* NRS 2004/3a variant 1 or *B. sphaericus* CCM 2177 were incubated for 2 d in buffered (pH 7.2) 10 mM CdCl₂ solution. Drops of this suspension were then mounted on Ni TEM grids and placed in a vacuum dessicator containing acidified Na₂S solution. Hydrogen sulphide gas was allowed to diffuse into the droplets over a period of 3 h, during which the grids were removed sequentially at regular time intervals, blotted on filter paper, washed with distilled water and left to dry in the air.

Received 11 March; accepted 8 July 1997.

1. Nanostructured Materials (spec. iss. eds Bein, T. & Stucky, G. D.) *Chem. Mater.* **8**, 1569–2194 (1996).
2. Mann, S. & Ozin, G. A. Synthesis of inorganic materials with complex form. *Nature* **382**, 313–318 (1996).
3. Chan, Y. N. C., Schrock, R. R. & Cohen, R. E. Synthesis of single silver nanoclusters within spherical microdomains in block copolymer films. *J. Am. Chem. Soc.* **114**, 7295–7296 (1992).
4. Spatz, J. P., Roescher, A. & Möller, A. Gold nanoparticles in micellar poly(styrene)- β -poly(ethylene oxide) films; size and interparticle distance control in monodisperse films. *Adv. Mater.* **8**, 337–334 (1996).
5. Burkett, S. L. & Mann, S. Spatial-organization and patterning of gold nanoparticles on self assembled biolipid tubular templates. *Chem. Commun.* 321–322 (1996).
6. Braun, P. V., Osener, P. & Stupp, S. I. Semiconducting superlattices templated by molecular assemblies. *Nature* **380**, 325–328 (1996).
7. Murray, C. B., Kagan, C. R. & Bawendi, M. G. Self-organization of CdSe nanocrystals into 3-dimensional quantum dot super lattices. *Science* **270**, 1335–1338 (1995).
8. Vossmeier, T. et al. A double diamond superlattice built up of Cd₁₇S₉(SCH₂CH₂OH)₍₂₆₎ clusters. *Science* **267**, 1476–1479 (1995).
9. Motte, L., Billoudet, F., Lacaze, E. & Pileni, M.-P. Self-organization of size-selected nanoparticles into three dimensional superlattices. *Adv. Mater.* **8**, 1018–1020 (1996).
10. Colvin, V. L., Goldstein, A. N. & Alivisatos, A. P. Semiconductor nanocrystals covalently bound to metal surfaces with self-assembled monolayers. *J. Am. Chem. Soc.* **114**, 5221–5230 (1992).
11. Dabbousi, B. O., Murray, C. B., Rubner, M. F. & Bawendi, M. G. Langmuir-Blodgett manipulation of size selected CdSe nanocrystals. *Chem. Mater.* **6**, 216–219 (1994).
12. Whetten, R. L. et al. Nanocrystal gold molecules. *Adv. Mater.* **8**, 428–433 (1996).
13. Kimizuka, N. & Kunitake, T. Organic 2-dimensional templates for the fabrication of inorganic nanostructures. *Adv. Mater.* **8**, 89–91 (1996).
14. Brust, M., Bethall, D., Schiffrin, D. J. & Kiely, C. J. Novel gold dithiol nanonetworks with non-metallic electronic properties. *Adv. Mater.* **7**, 795–797 (1995).
15. Andres, R. P. et al. Self-assembly of a 2-dimensional superlattice of molecularly linked metal clusters. *Science* **273**, 1690–1693 (1996).
16. Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382**, 607–609 (1996).
17. Alivisatos, A. P. et al. Organization of nanocrystal molecules using DNA. *Nature* **382**, 609–611 (1996).
18. Heath, J. R. et al. Spatially confined chemistry: fabrication of Ge quantum dot arrays. *J. Phys. Chem.* **100**, 3144–3149 (1996).
19. Sleytr, U. B., Messner, P., Pum, D. & Sara, M. (eds) *Crystalline Bacterial Cell Surface Proteins* (Landes Co., Austin, & Academic, San Diego, 1996).
20. Beveridge, T. J. Bacterial S-layers. *Curr. Opin. Struct. Biol.* **4**, 204–212 (1994).
21. Sara, M. & Egelseer, E. in *Crystalline Bacterial Cell Surface Proteins* (eds Sleytr, U. B., Messner, P., Pum, D. & Sara M.) 103–131 (Landes Co., Austin, & Academic, San Diego, 1996).
22. Schultze-Lam, S. & Beveridge, T. J. Nucleation of celestite and strontianite on a cyanobacterial S-layer. *Appl. Environ. Microbiol.* **60**, 447–453 (1994).
23. Lowenstam, H. A. in *Mechanisms and Phylogeny of Mineralization in Biological Systems* (eds Suga, S. & Nakara, H.) 3–8 (Springer, Tokyo, 1991).
24. Sara, M., Pum, D., Kuepfer, S., Messner, P. & Sleytr, U. B. Isolation of two physiologically induced variant strains of *Bacillus stearothermophilus* NRS 2004/3A and characterisation of their lattices. *J. Bacteriol.* **176**, 848–860 (1994).
25. Messner, P., Pum, D. & Sleytr, U. B. Characterisation of the ultrastructure and self assembly of the S-layer of *Bacillus stearothermophilus* strain NRS 2004/3A. *J. Ultrastruct. Mol. Struct. Res.* **97**, 73–88 (1986).
26. Sleytr, U. B. & Sara, M. Bacterial and archaeal S-layer proteins: Structure-function relationships and their biotechnological applications. *Trends Biotechnol.* **15**, 20–26 (1997).
27. Douglas, K., Devaud, G. & Clark, N. A. Transfer of biologically derived nanometer scale patterns to

smooth substrates. *Science* **257**, 642–644 (1992).

28. Tucker, J. R. Complementary digital logic based on the coulomb blockade. *J. Appl. Phys.* **71**, 4399–4402 (1992).
29. Pum, D., Sara, M. & Sleytr, U. B. Structure, surface charge and self assembly of the S-layer lattice from *Bacillus coagulans* E38-55. *J. Bacteriol.* **171**, 5296–5303 (1989).
30. Pum, D. & Sleytr, U. B. Large scale reconstitution of crystalline S-layer proteins at the air–water interface and on lipid films. *Thin Solid Films* **244**, 882–886 (1994).

Acknowledgements. W.S. thanks the University of Bath for a postgraduate studentship. This work was supported in part by the Austrian Science Foundation, Project S7204 and S7205, the Austrian Ministry of Science and Transportation, and the Austrian National Bank.

Correspondence and requests for materials should be addressed to S.M. (e-mail: s.mann@bath.ac.uk) or U.B.S. (e-mail sleytr@edv1.boku.ac.at).

Convergent total synthesis of a tumour-associated mucin motif

Dalibor Sames*, Xiao-Tao Chen†
& Samuel J. Danishefsky*†

* Laboratory for Bioorganic Chemistry, Sloan-Kettering Institute for Cancer Research, 1275 York Avenue, New York, New York 10021, USA

† Department of Chemistry, Columbia University, New York, New York 10027, USA

Synthetic glycoconjugates that mimic cell-surface tumour antigens (glycolipids or glycoproteins with unusual carbohydrate structural motifs) have been shown to trigger humoral responses in murine and human immune systems^{1–3}. This raises the exciting possibility of inducing active immunity with fully synthetic carbohydrate vaccines, particularly if vaccine compounds can be synthesized that resemble the surface environment of transformed cells even more closely. Glycopeptides seem particularly suitable for this purpose. In contrast to most glycolipids and the carbohydrates themselves, glycopeptides bind to major histocompatibility complex molecules, and, in favourable cases, can stimulate T cells and lead to the expression of receptors that recognize the carbohydrate part of a glycopeptide with high specificity^{4–8}. The preparation of glycopeptides and glycoproteins remains, however, a difficult challenge^{9–12}; earlier synthesis methods have been inefficient, and established cloning approaches that allow engineering of global glycopatterns produce only heterogeneous glycoproteins¹³. Here we report an efficient strategy of the synthesis of tumour-associated mucin glycopeptides with clustered trisaccharide glycodomains corresponding to the (2,6)-sialyl T antigen. Our approach involves construction of the complete glycodomain in the first stage, followed by convergent coupling to amino acid residues and subsequent incorporation of the glycosyl amino acid units into a peptide chain. This general strategy allows the assembly of molecules in which selected glycoforms can be incorporated at any desired position of the peptide chain. The resultant fully synthetic O-linked glycopeptide clusters are the closest homogeneous mimics of cell-surface mucins at present available, and so are promising compounds for the development of anticancer vaccines.

To demonstrate the logic of our approach, we describe here the synthesis of the mucin (2,6)-sialyl T antigen (ST antigen), which is an example of the 'glycophorin family' of α -O-linked glycopeptides (Fig. 1). Mucin glycoproteins are an important structural class of cell-surface molecules¹⁴. Mucins containing aberrant α -O-glycosidation patterns which are characterized by clustered arrangements of carbohydrates α -O-linked to serine and threonine residues are common markers of epithelial tumours (for example, prostate and breast carcinomas) and some tumours of blood cells (Fig. 1)¹⁵. The

ST antigen is selectively expressed on myelogenous leukaemia cells^{16,17}. Because of the large size of mucin glycoproteins, we selected the most accessible amino-terminal fragments of mucin CD43 containing clustered carbohydrate epitopes^{14,18}. The pentapeptide **1** (Fig. 1b) served as our entry point.

Chemical synthesis of α -O-linked glycopeptides has been a challenge for many years^{12,19}. Nearly all approaches have incorporated the amino acid (serine or threonine) at a monosaccharide stage. This construction would be followed by elaboration of the peptidyl and carbohydrate domains in piecemeal fashions^{20–24}.

Given the scope of the problem, the progress which has been achieved so far is remarkable. To test our concepts at the biological and immunological levels, it would be necessary to produce substantial quantities of homogeneous glycopeptide for methodical evaluation. This need prompted us to explore new ideas to achieve convergency in the assembly of complex, O-linked clustered glycopeptides. The convergent approach used here represents a significant advance in terms of speed and scope.

Our starting point was the readily available glycal **2** (Fig. 2). Oxidation of this compound with dimethyldioxirane and subse-

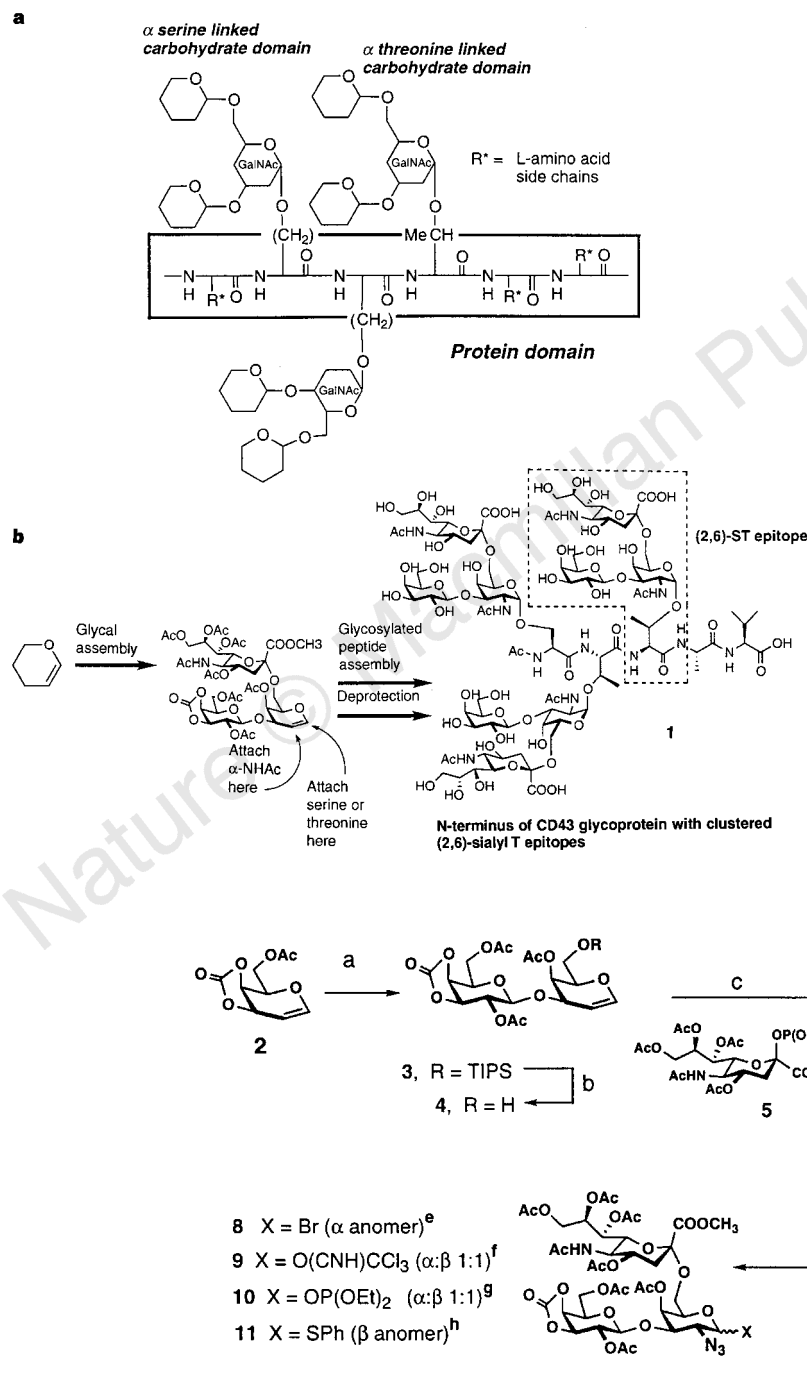


Figure 1 Illustration of α -O-linked glycopeptides and the synthetic strategy reported here. **a**, O-Linked glycoproteins as found in mucins; GalNAc, *N*-acetylgalactosamine. **b**, General synthetic strategy to mucin glycopeptides; Ac, acetyl.

Figure 2 Assembly of trisaccharide glycal and donors. **a**, (1) DMDO, CH₂Cl₂; (2) 6-O-TIPS-galactal, ZnCl₂, -78 to 0°C; (3) Ac₂O, Et₃N, DMAP, 75%; **b**, TBAF/AcOH/THF; 80%; **c**, **5** (1.3 equiv.), TMSOTf (0.1 equiv.), THF:toluene 1:1, -60 to -45°C, 84%, α : β 4:1; **d**, NaN₃, CAN, CH₃CN, -15°C, 60%; **e**, LiBr, CH₃CN, 75%; **f**, (1) PhSH, *iso*-Pr₂NEt, CH₃CN, 82%; (2) CCl₃CN, K₂CO₃, CH₂Cl₂, 80%; **g**, (1) PhSH, *iso*-Pr₂NEt;

(2) ClP(OEt)₂, *iso*-Pr₂NEt, THF, (labile compound, ~72% for two steps); **h**, (1) LiBr, CH₃CN, 75%; (2) LiSPh, THF, 0°C, 70%. Ac, acetyl; N₃, azide; TIPS, triisopropyl silyl; Et, ethyl; Ph, phenyl; DMDO, dimethyldioxirane; TMSOTf, trimethylsilyl trifluoromethanesulphonate; DMAP, dimethylaminopyridine; TBAF, tetrabutylammonium fluoride; THF, tetrahydrofuran; CAN, cerium ammonium nitrate.

quent coupling of the resultant epoxide with 6-*O*-TIPS-galactal (where TIPS is tri-isopropyl silyl) was promoted by ZnCl₂ in the usual way³. Acetylation of the crude product yielded disaccharide **3** in high yield and high stereoselectivity. Removal of the TIPS protecting group under mild conditions allows the attachment of sialic acid to acceptor **4**. The use of sialyl phosphite **5** as the donor, under promotion of catalytic amounts of trimethylsilyl trifluoromethanesulphonate (TMSOTf), consistently provided high yields (80–85%) of a 4:1 mixture of products^{25,26}. Thus, the advanced glycal **6** ('2,6-ST glycal') is available in four steps with high efficiency.

The trisaccharide glycal **6** was submitted to azidonitration as shown (Fig. 2). Compound **7**, thus obtained in 60% yield, lent itself to conversion to a variety of donor constructs (see **8–11**). For instance, α -bromide **8** can be used as a donor directly or could be converted to β -phenylthioglycoside **11** with lithium thiophenoxide in a stereoselective manner. Alternatively, the mixture of nitrates **7** was hydrolysed and the resulting hemiacetal converted to a 1:1 mixture of α : β trichloroacetimidates **9** and diethylphosphites **10** in high yields (Fig. 2)^{27,28}.

The availability of various donor types (**8–11**) enables us to investigate extensively the direct coupling of (2,6)-ST trisaccharide to the benzyl ester of *N*-Fmoc-protected L-serine and L-threonine (where Fmoc is fluorenylmethoxycarbonyl). The results are summarized in Table 1. With Fmoc-protected L-threonine as the acceptor, all of the donors afforded the α -*O*-glycosyl threonine system with high stereoselectivity. The outcome of the coupling reactions with similarly protected L-serine acceptors was dependent on the character of the donor and on the reaction conditions. With donor **10** the ratio of desired α -product to undesired β -glycoside was $\sim$ 30:1. In all cases, the desired α -anomer **12** was the main product.

Previous attempts to fashion a convergent coupling of trisaccharide donors to serine have failed, yielding the β -isomer as the main

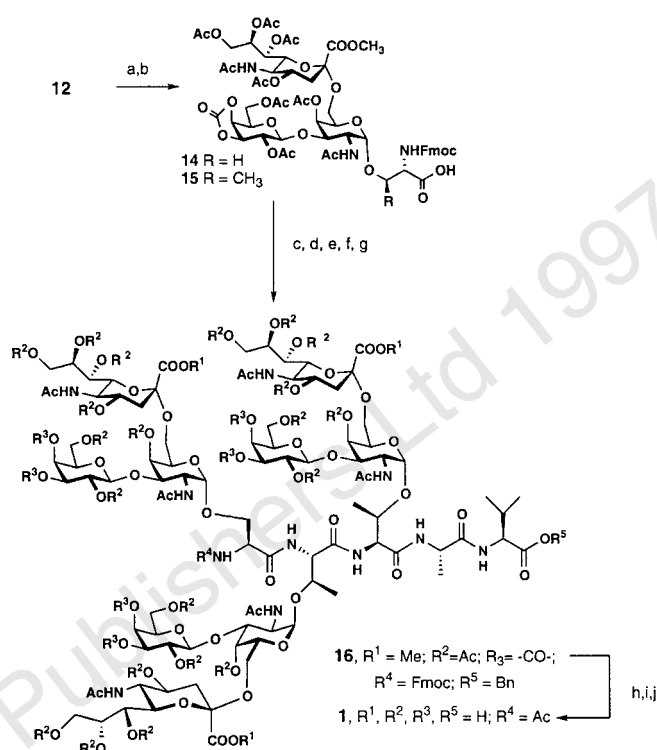


Figure 3 Glycopeptide assembly and deblocking of protecting groups. a, CH₃COSH, 78%; b, H₂/10% Pd-C, MeOH, H₂O, quant; c, H₂N-Ala-Val-OBn, IIDQ, CH₂Cl₂, 85%; d, KF, DMF, 18-crown-6, 95%; e, **15**, IIDQ, 87%; f, KF, DMF, 18-crown-6, 93%; g, **14**, IIDQ, 90%; h, (1) KF, DMF, 18-crown-6, (2) Ac₂O, CH₂Cl₂; i, H₂/10% Pd-C (20 wt%), 2.4 μ M in MeOH/H₂O (15:1), RT; j, 5 μ M in MeOH, 1 M NaOH, pH $\sim$ 10–10.5, 80%.

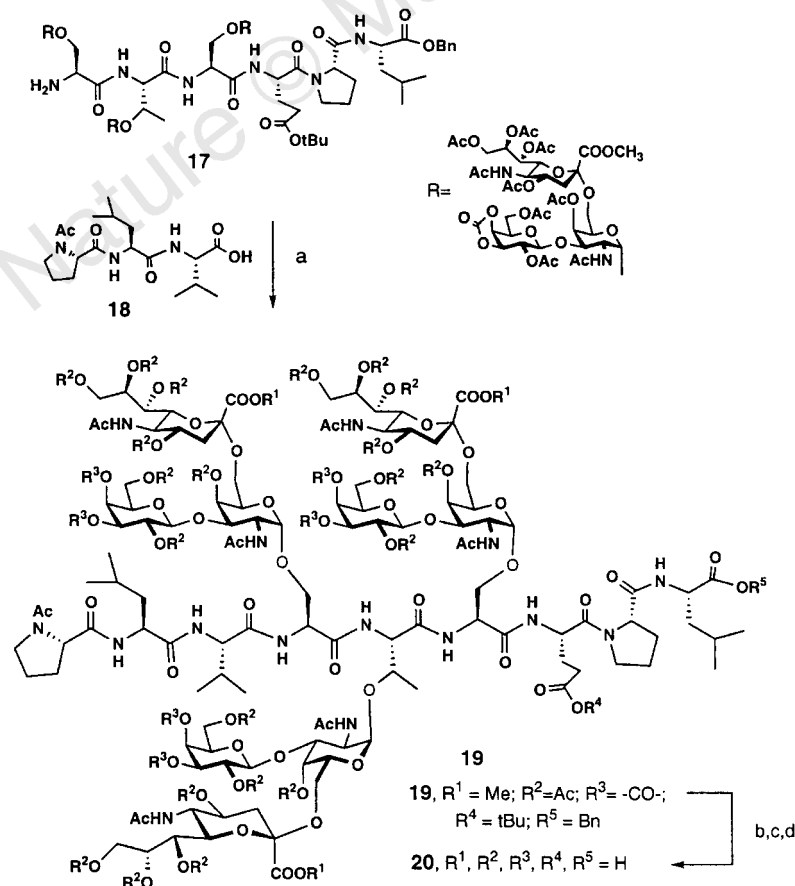
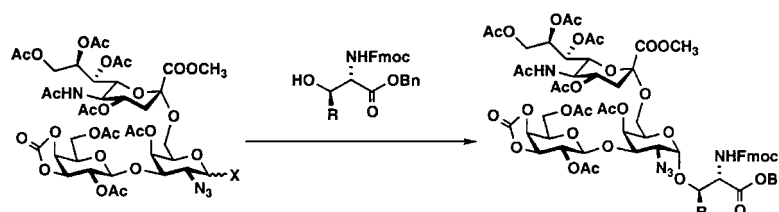


Figure 4 Fragment assembly of glycoprotein **20**. a, IIDQ, CH₂Cl₂, RT, 80%; b, H₂/Pd-C, MeOH, H₂O, 95%; c, CF₃COOH, CH₂Cl₂; d, NaOH, H₂O, MeOH, 70% in two steps. tBu, *t*-butyl.



Fmoc, fluorenylmethoxycarbonyl; Bn, benzyl. Protocol A: donor **8** (1.05 equiv, 0.16 M in CH_2Cl_2) was added over 30 min to acceptor (0.09 M in CH_2Cl_2), AgClO_4 (1.5–2 equiv.) and 4-Å molecular sieves (1.7 g mmol^{-1}) at room temperature. Protocol B: donor **9** and **10** (1 equiv, 0.07 M in THF) and acceptor (1.5 equiv.) and 4-Å molecular sieves (1.8 g mmol^{-1}) were cooled to -30°C and then treated with $\text{BF}_3\cdot\text{OEt}_2$ (0.5 equiv.).

We decided to enter the glycopeptide assembly phase with building units **14** and **15** (Fig. 3), thereby reducing the number of required chemical operations to be performed on the final glycopeptide. In fact, compounds **14** and **15** were obtained in two steps from **12** and **13**, respectively. The azide functionality was transformed directly to *N*-acetyl groups by the action of CH_3COSH in 78–80% yield and the benzyl ester was removed quantitatively by hydrogenolysis (Fig. 3).

The orthogonal exposure of both N and C termini provided an opportunity for further extension of this glycopeptide construct by fragment joining. To demonstrate the viability of this approach, we prepared a nonapeptide with ST triad **19** by means of coupling tripeptide **18** to hexapeptide **17**. The above-mentioned deprotection protocol provided nonapeptide mucin model **20**, in which the O-glycosylated serine-threonine triad had been incorporated in the middle of the peptide (see Fig. 4).

Received July 1978; revised November 1978

1. Ragupathi, G. *et al.* Immunization of mice with a fully synthetic globo-H antigen results in antibodies against human cancer cells: A combined chemical-immunological approach to the fashioning of an anticancer vaccine *Angew. Chem. Int. Edn Engl.* **36**, 125–128 (1997).

31. Jiang, J., Li, W.-R. & Joullie, M. Selective removal of fluorenylmethoxycarbonyl (Fmoc) groups under mild conditions. *Synth. Commun.* **24**, 187–195 (1994).

Acknowledgements. We thank G. Sukenick for helping with NMR and mass spectrometry. This work was supported by the NIH; D.S. thanks the Irvington Institute for Immunological Research and M.R. Bloomberg for a postdoctoral fellowship.

Correspondence and requests for materials should be addressed to S.J.D. (e-mail: s-danishsky@ski.mskcc.org).

Magma mixing as a source for Pinatubo sulphur

Victor Kress

Department of Geological Sciences, University of Washington, Box 351310, Seattle, Washington 98195-1310, USA

On 15 June 1991, a huge plinian eruption at Mount Pinatubo discharged 3.7–5.3 km³ of pyroclastic material¹, along with a minimum of 17 megatonnes of SO₂ gas^{2,3}. This represents the largest stratospheric SO₂ cloud ever measured, and the SO₂ generated in this eruption is believed to have had a significant effect on global climate^{2,4} and the ozone layer⁴ for several years after the event. The source for this massive amount of SO₂ aerosols remains controversial. Here I present thermodynamic arguments which suggest that the source of the SO₂, along with the trigger for the eruption itself, can be attributed to redox reactions accompanying the injection of a reduced sulphide-saturated basaltic magma into an oxidized sulphate-saturated dacitic melt. The proposed mixing event would drive most sulphur out of both dacitic and basaltic liquids and would drive both anhydrite and iron-rich sulphide liquid outside their stability field, thus purging sulphur from all major non-volatile sulphur-bearing phases in the mixed volume. Similar eruptions are possible any time that an oxidized sulphate-saturated magma interacts with a reduced sulphide-saturated magma, and this mechanism may therefore be relevant to recent volcanic activity at Popocatepetl Volcano in Mexico.

Pyroclastic products of the 15 June Pinatubo eruption consist primarily of a phenocryst-rich dacitic pumice^{5,6}. This pumice contains Cu-rich sulphide globules, along with phenocrysts of primary anhydrite^{5,7}. Between 7 June and the climactic eruption of 15 June, Mt Pinatubo erupted ~0.1 km³ of andesitic lava and scoria (C. G. Newhall, personal communication).

Both dacite matrix glass and melt inclusions in phenocryst phases contain ~90 p.p.m. dissolved sulphur³. This suggests that little of the sulphur dissolved in the dacitic liquid was liberated to the volatile phase at the time of eruption⁸. Anhydrite and Cu-rich sulphides represent the other significant non-volatile reservoirs of sulphur in the dacite. Both appear to be relatively stable phases in the dacite^{7,9}. Both anhydrite and sulphide liquid are observed as inclusions in phenocryst phases in the dacite, suggesting that these phases were saturated well before eruption.

Pinatubo dacite is highly oxidized. A recent critical assessment¹⁰ suggests that pre-eruption oxygen fugacity (f_{O_2}) for the Pinatubo dacite was probably 2.3–2.9 log₁₀ units higher than that defined by the fayalite + magnetite + quartz assemblage at the temperature and pressure in question (that is, $\Delta_{FMQ} = 2.3$ –2.9). The presence of anhydrite inclusions in phenocrysts confirms that the oxidized character of these dacites was established well before eruption.

The andesitic precursor to the climactic 15 June eruption was volumetrically minor relative to the dacite, but this belies the significance of these andesites in the Pinatubo story. Though the bulk rock composition of this material is andesitic, matrix glass compositions are variable on the thin-section scale. Furthermore, Pinatubo andesites contain phenocryst assemblages which are

significantly out of equilibrium^{11,12}. Pallister *et al.*^{11,12} suggested that these andesites are a product of mixing between the dacitic magma described above and a basaltic liquid. This basaltic liquid is represented in the eruption products in the form of isolated millimetre- to decimetre-scale blebs in the hybrid andesite¹². Iron-rich sulphides are present as inclusions in basalt phenocrysts.

The original oxidation state of the basalt is difficult to establish with certainty. Fe–Ti oxides in the basalt blebs have been reset from some more reduced state by interaction with the dacite¹³. It is probably safe to assume $\Delta_{FMQ} = 0$, as this is on the oxidized side of values typically observed in basaltic magmas¹⁴.

Experiments^{15–18} in silicate liquids indicate that sulphur solubility passes through a minimum at $\Delta_{FMQ} \approx 0.7$ (Fig. 1). Mixing of a reduced sulphide-saturated basalt with an oxidized sulphate-saturated dacite would drive both magmas towards this solubility minimum. Oxygen diffusion rates in silicate liquids¹⁹, along with the observed millimetre length scales of cation inhomogeneity in andesite glass, suggest that redox equilibration of the andesite liquid can be expected within ~90 min of the mixing event. Sulphide should reach equilibrium about four times faster than the silicate²⁰. Assuming $\Delta_{FMQ} = 2.6$ for the dacite and zero for the basalt, a mixture of 63.5% dacite and 36.5% basalt¹² can be predicted²¹ to have a Δ_{FMQ} value of 0.8–0.9. It has been observed experimentally^{17,18} that, to a first approximation, $\log_{10}(X_S) - 1/2 \log_{10}(f_{S_2}) - 3/2 \log_{10}(f_{O_2}) = C1$ below this minimum and $\log_{10}(X_S) - 1/2 \log_{10}(f_{S_2}) + 1/2 \log_{10}(f_{O_2}) = C2$ above this minimum, where X_S is the mole fraction of sulphur in the melt. C1 and C2 are constants which can be estimated from the initial sulphur contents of liquids in the basalt and dacite, respectively. Thus, neglecting compositional and thermal contributions and assuming constant f_{S_2} , mixing of these two magmas will liberate 1,124 p.p.m. out of 1,800 p.p.m. sulphur in the basalt and 89.8 p.p.m. out of 90 p.p.m. sulphur in the dacite. The assumption of constant f_{S_2} is not strictly correct, but the buffering effect of the water-dominated vapour phase²² will limit the increase in sulphur fugacity expected to accompany this process.

Experiments²⁰ on silicate-saturated sulphide liquids in the system

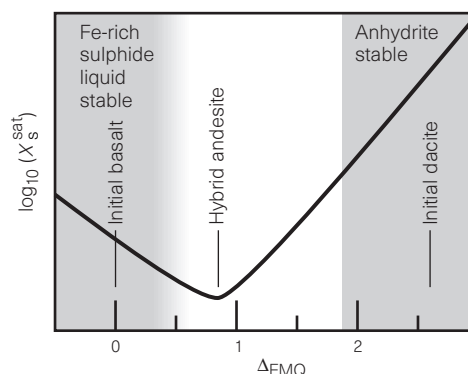


Figure 1 Schematic diagram of the solubility of sulphur in silicate magma, where X_S^{sat} is the mole fraction of S at vapour saturation, as a function of oxygen fugacity in Δ_{FMQ} (see text). Also shown are the stability fields of Fe-rich sulphide liquids and anhydrite. Sulphide liquids in basalts and anhydrite phenocrysts in dacites indicate that both liquids had a high initial sulphur fugacity. Coexisting magnetite and pyrrhotite in the dacite suggest a sulphur fugacity (f_{S_2}) about 2.5 log₁₀ units above that defined by the fayalite + magnetite + quartz + pyrrhotite assemblage ($\Delta_{FMQ} \approx 2.5$)³¹. Increasing f_{S_2} by one log₁₀ unit will increase the sulphide stability field by one log₁₀ unit in f_{O_2} (ref. 20) and will expand the anhydrite stability field by about 1/3 log₁₀ unit in f_{O_2} . Lack of an explicit pressure correction is equivalent to assuming that the volume of the sulphide liquid breakdown reaction is roughly equal to that of the pyrrhotite breakdown reaction on an equal S₂ basis. This is a reasonable first approximation in the absence of experimental data. $\Delta_{FMQ} = 2.5$ corresponds to $\log_{10}(f_{S_2}) = -1.5$ at atmospheric pressure.

O–S–Fe–Si indicate that sulphide liquids are not stable with a silica-saturated silicate liquid at low sulphur fugacities or high oxygen fugacities. For example, at $\log_{10}(f_{S_2}) = -1.5$ and $1,250^\circ\text{C}$, sulphide liquids in the system O–S–Fe cannot coexist with a silica-saturated silicate liquid at Δ_{FMQ} above 0.5 (Fig. 1). Oxidation of a reduced silicate liquid saturated with sulphide liquid will drive iron from the latter liquid into the former. This will increase sulphur fugacity to vapour saturation, driving sulphur into the vapor phase as SO_2 . Ni and Cu expand the stability field of sulphide liquids substantially. Oxidation of a sulphide with trace Ni and Cu will decompose most of the sulphide in the manner described above, leaving a trace amount of residual Ni–Cu-rich sulphide liquid. Thus, Fe-rich sulphide liquids are indicators of reducing conditions. This is why there are no Fe-rich sulphides in Pinatubo dacites. Sulphide inclusions in basalt phenocrysts have¹³ Fe/(Ni + Cu) on the order of 50, further supporting the assertion that these magmas were originally rather reduced.

Anhydrite is only stable at^{23,24} $\Delta_{\text{FMQ}} > 1.8$ at 800°C and $\Delta_{\text{FMQ}} \approx 2.5$ (Fig. 1). Reduction of an anhydrite-bearing magma below this f_{O_2} will initiate breakdown of the anhydrite, driving CaO into the silicate liquid and SO_2 into the vapour.

The proposed mixing event would drive the oxidation state of the dacite below the stability field of anhydrite. At the same time, it would drive the basalt above the stability field of its Fe-rich sulphide liquids. Thus, the proposed magma mixing event can be expected to purge most of the sulphur from all the main non-volatile sulphur-bearing phases in the mixed volume. Assuming 0.01% sulphide liquid in the basalt^{9,13} and using sample P4 from Fournelle *et al.*⁵ to calculate the anhydrite contribution in the dacite, I estimate that 17 megatonnes (Mt) of SO_2 aerosol can be generated by mixing 1.2 km^3 of sulphide saturated basaltic magma into the dacitic magma chamber. The relative contribution of each of the sulphur sources to this 17 Mt total is shown in Fig. 2. The proposed 1.2 km^3 of basalt involved in the mixing event probably represents a small portion of the total injected basalt. Geophysical evidence²⁵ suggests that the combined magma chamber beneath Pinatubo may be as large as $40\text{--}90\text{ km}^3$.

Snyder and Tait²⁶ propose a mechanism which explains how a

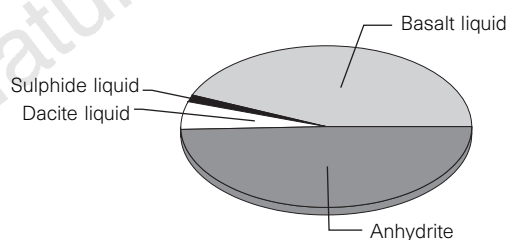


Figure 2 Relative contributions of the main sulphur-bearing phases to the $>17\text{ Mt}$ total SO_2 discharged in the 15 June Pinatubo eruption, based on assumptions given in the text. The original sulphide content of the basalt is not well constrained, but the final result is not very sensitive to variations in this parameter. Sulphide liquid was assumed to make up 0.01% of the basalt based on the fact that this is a typical value for oceanic basalts. Increasing the proportion of sulphide in the basalt to 0.1% will increase the sulphide contribution to 13% and will decrease the volume of basalt required in the mixed volume to 1.1 km^3 . Lacking direct measurements of the sulphur content of the basaltic liquid, we adopt a value of 0.18 wt%. This is towards the lower end of the range of sulphur solubilities observed in basaltic melt inclusions from subduction-related magmas³² but is towards the upper end of experimentally derived values^{33–35}. Unreacted and incompletely reacted anhydrite crystals in andesite^{11,12} suggest that the anhydrite contribution cited above is probably an overestimate. Assuming zero contribution from anhydrite, 2.5 km^3 of basalt is required to produce the observed SO_2 output at Pinatubo.

hybrid magma resulting from a basalt injected from below into a dacitic magma chamber could have erupted before the main volume of dacite. They suggest that the injection of basalt generated a thermal perturbation in the overlying dacite, inducing the formation of a convective plume. Viscous coupling between the dacite and the basalt entrained chilled basaltic magma in the rising superheated dacite plume. This served both to mix the two magmas, and to bring the hybrid magma to the top of the magma chamber. Once there, the hybrid magma would be available for eruption before the main body of dacite.

Mixing between this reduced sulphide-saturated basalt and the oxidized sulphate-saturated dacite will generate large amounts of SO_2 vapour. Generation of a vapour phase by magma mixing will decrease the density of the magma locally, increasing the rate of convection. This will increase the rate of entrainment of the basalt, leading to further SO_2 generation. Assimilation of dacite will also reduce the density of the hybrid magma relative to the original basalt. Dilution of the vapour phase with SO_2 in the rising plume will scavenge more H_2O out of the surrounding magma, further increasing the volume of the vapour phase and accelerating the process. This runaway mechanism could liberate large amounts of sulphur-rich vapour relatively quickly, and may have served as the trigger for the climactic eruption of 15 June. Pallister *et al.*¹¹ also suggested that magma mixing may have initiated eruption, though their proposed triggering mechanism is quite different from the one proposed here.

The model presented here is not the only feasible explanation for the generation of Pinatubo aerosols. It is widely acknowledged that Pinatubo dacites were vapour-saturated well before eruption²². Using Wallace and Gerlach's²² estimate of the vapour-phase composition and an estimated magma chamber pressure of $\sim 220\text{ MPa}$ (ref. 27); 17 Mt of SO_2 corresponds to $\sim 0.32\text{ km}^3$ of vapour phase. This represents 6–9% of the total erupted dacite volume. If one is to call on this volatile phase alone as the primary source for Pinatubo SO_2 , one must provide a mechanism for concentration and storage of this volatile phase before eruption. Volatiles could have been derived by volatile fluxing from an external source, or could have been derived from the magma body itself by decompression, crystallization and/or concentration from a larger body. Any of these processes would require extended periods of time. For such a model to be viable, the flux of volatiles from an external source plus the rate of internal generation of volatiles must exceed the flux of volatiles out of the magma body. Furthermore, any external source must be as oxidized as the dacites. This probably rules out the basalt as an external source. Regardless of the source of the vapour, it is not clear that a magma chamber could sustain 6–9% volatile content for extended periods without erupting or degassing. This question may be a fruitful topic for future research.

The magma-mixing model for the generation of Pinatubo sulphur presented here does not require extended volatile accumulation and storage. The model explains the sulphur-rich nature of Pinatubo aerosols, ties in the associated precursor andesite eruption, and provides a trigger mechanism for the climatic 15 June eruption. Finally, the model presented here predicts conditions which could lead to similar sulphur-rich eruptions at other localities. The magma mixing and vapour-accumulation hypotheses are not mutually exclusive. Both processes probably contributed to the June 1991 eruptive sequence at Pinatubo.

The magma mixing/sulphur redox mechanism proposed for the 15 June Pinatubo eruption can be expected to apply any time an oxidized sulphate-saturated magma is injected with a reduced sulphate-saturated magma. Anhydrite-saturated magmas are now recognized to be relatively common in nature. For the most part, eruption of these anhydrite-saturated magmas takes place without the apparent influence of a second magma. But if such a magma chamber is injected with a reduced sulphide-saturated magma, the

resulting eruption can be expected to be more violent and discharge more SO₂ into the atmosphere than would be observed without the mixing event.

Juvenile products of recent activity at Volcan Popocatepetl, Mexico, exhibit both primary igneous anhydrite²⁸ and evidence for magma mixing²⁹. Eruptive events have been relatively small so far. Nevertheless, Popocatepetl has been generating SO₂ at sustained high rates³⁰ which are difficult to account for from conventional sources. It is far too early to draw conclusions, but evidence hints that the same sort of process proposed for the June 1991 activity at Pinatubo may be at work today at Popocatepetl. □

Received 29 January; accepted 11 August 1997.

1. Scott, W. E. et al. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 545–570 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
2. Bluth, G. J. S., Doiron, S. D., Schnetzler, C. C. & Krueger, A. J. Global tracking of the SO₂ clouds from the June, 1991 Mount Pinatubo eruptions. *Geophys. Res. Lett.* **19**, 151–154 (1992).
3. Gerlach, T. M., Westrich, H. R. & Symonds, R. B. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 415–434 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
4. Brasseur, G. & Granier, C. Mount Pinatubo aerosols, chlorofluorocarbons, and ozone depletion. *Science* **257**, 1239–1242 (1992).
5. Fournelle, J., Carmody, R. & Daag, A. S. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 845–864 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
6. Imai, A., Listanco, E. L. & Fujii, T. Petrologic and sulfur isotopic significance of highly oxidized and sulfur-rich magma of Mt. Pinatubo, Philippines. *Geology* **21**, 699–702 (1993).
7. Bernard, A., Demaffie, D., Mattioli, N. & Punongbayan, R. S. Anhydrite-bearing pumices from Mount Pinatubo: further evidence for the existence of sulphur-rich silicic magmas. *Nature* **354**, 139–140 (1991).
8. Westrich, H. R. & Gerlach, T. M. Magmatic gas source for the stratospheric SO₂ cloud from the June 15, 1991, eruption of Mount Pinatubo. *Geology* **20**, 867–870 (1992).
9. Hattori, K. High-sulfur magma, a product of fluid discharge from underlying mafic magma: Evidence from Mount Pinatubo, Philippines. *Geology* **21**, 1083–1086 (1993).
10. Evans, B. W. & Scaillet, B. The redox state of Pinatubo dacite and the ilmenite-hematite solvus. *Am. Mineral.* **82**, 625–629 (1997).
11. Pallister, J. S., Hoblitt, R. P. & Reyes, A. G. A basalt trigger for the 1991 eruptions of Pinatubo volcano? *Nature* **356**, 426–428 (1992).
12. Pallister, J. S., Hoblitt, R. P., Meeker, G. P., Knight, R. J. & Siems, D. F. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 687–732 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
13. Hattori, K. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 807–824 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
14. Carmichael, I. S. E. The redox states of basic and silicic magmas: a reflection of their source regions? *Contrib. Mineral. Petrol.* **106**, 129–141 (1991).
15. Carroll, M. R. & Rutherford, M. J. Sulfide and sulfate saturation in hydrous silicate melts. *J. Geophys. Res.* **90**, C601–C612 (1985).
16. Katsura, T. & Nagashima, S. Solubility of sulfur in some magmas at 1 atmosphere. *Geochim. Cosmochim. Acta* **38**, 517–531 (1974).
17. Fincham, C. J. B. & Richardson, F. D. The behavior of sulphur in silicate and aluminate melts. *Proc. R. Soc. Lond. A* **223**, 40–62 (1954).
18. Nagashima, S. & Katsura, T. The solubility of sulfur in Na₂O-SiO₂ melts under various oxygen partial pressures at 1100°C, 1200°C and 1300°C. *Bull. Chem. Soc. Jpn.* **46**, 3099–3103 (1973).
19. Wendlandt, R. F. Oxygen diffusion in basalt and andesite melts: experimental results and discussion of chemical versus tracer diffusion. *Contrib. Mineral. Petrol.* **108**, 463–471 (1991).
20. Kress, V. C. Thermochemistry of sulfide liquids I: The system O-S-Fe at 1 bar. *Contrib. Mineral. Petrol.* **127**, 176–186 (1997).
21. Kress, V. C. & Carmichael, I. S. E. The compressibility of silicate liquids containing Fe₂O₃ and the effect of composition, temperature, oxygen fugacity and pressure on their redox states. *Contrib. Mineral. Petrol.* **108**, 82–92 (1991).
22. Wallace, P. J. & Gerlach, T. M. Magmatic vapor source for sulfur dioxide released during volcanic eruptions: Evidence from Mount Pinatubo. *Science* **265**, 497–499 (1994).
23. Luhr, J. F. Experimental phase relations of water- and sulfur-saturated arc magmas and the 1982 eruptions of El Chichón Volcano. *J. Petrol.* **5**, 1071–1114 (1990).
24. Carroll, M. R. & Rutherford, M. J. The stability of igneous anhydrite: experimental results and implications for sulfur behavior in the 1982 El Chichón trachyandesite and other evolved magmas. *J. Petrol.* **28**, 781–801 (1987).
25. Mori, J., Eberhart-Phillips, D. & Harlow, D. H. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 371–382 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
26. Snyder, D. & Tait, S. Magma mixing by convective entrainment. *Nature* **379**, 529–531 (1996).
27. Rutherford, M. J. & Devine, J. D. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 751–766 (Philippine Inst. Volcanol. & Seismol./Univ. Washington Press, Quezon City/Seattle, 1996).
28. Del Pozzo, A. L. M., Espinosa, R. & Butron, M. A. Popocatepetl's 1994–1996 activity: eruption products. (abstr.) *Eos* **77**, F809 (1996).
29. Athanopoulou, P. & Laroque, A. C. L. Recent eruptions from Volcan Popocatepetl in context of past eruptions. (abstr.) *Eos* **77**, F809 (1996).
30. Delgado, H. & Cárdenas-González, L. (abstr.) Passive degassing at Volcan Popocatepetl (Mexico): 2.6 × 10⁶ tons of SO₂ released in 617 days of activity. *IAGCEI General Assembly Abstracts* 49 (Gobierno del Estado de Jalisco, Guadalajara, Jalisco, Mexico, 1997).
31. Whitney, J. A. Fugacities of sulfurous gases in pyrrhotite-bearing silicic magmas. *Am. Mineral.* **69**, 69–78 (1984).
32. Metrich, N. & Clocchiatti, R. Sulfur abundance and its speciation in oxidized alkaline melts. *Geochim. Cosmochim. Acta* **60**, 4151–4160 (1996).
33. Buchanan, D. L. & Nolan, J. Solubility of sulfur and sulfide immiscibility in synthetic tholeiitic melts and their relevance to Bushveld-Complex rocks. *Can. Mineral.* **17**, 483–494 (1979).

34. Peach, C. L. & Mathez, E. A. Sulfide melt-silicate melt distribution coefficients for nickel and iron and implications for the distribution of other chalcophile elements. *Geochim. Cosmochim. Acta* **57**, 3013–3021 (1993).
35. Houghton, D. R., Roeder, P. L. & Skinner, B. J. Solubility of sulfur in mafic magmas. *Econ. Geol.* **4**, 451–467 (1974).

Acknowledgements. I thank M. R. Carroll, B. W. Evans, M. S. Ghiorso, K. Hattori, N. Metrich, B. O. Mysen, B. K. Nelson, C. G. Newhall, J. S. Pallister, D. A. Snyder and P. J. Wallace for criticisms, suggestions, information and reviews.

Correspondence and requests for materials should be addressed to the author (e-mail: kress@u.washington.edu).

Fitness loss and germline mutations in barn swallows breeding in Chernobyl

Hans Ellegren*, Gabriella Lindgren*, Craig R. Primmer* & Anders Pape Møller†

* Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 597, S-751 24 Uppsala, Sweden

† Laboratoire d'Ecologie, CNRS URA 258, Université Pierre et Marie Curie, Bat. A, 7ème étage, 7, quai St Bernard, Case 237, F-75252 Paris Cedex 5, France

The severe nuclear accident at Chernobyl in 1986 resulted in the worst reported accidental exposure of radioactive material to free-living organisms¹. Short-term effects on human populations inhabiting polluted areas include increased incidence of thyroid cancer², infant leukaemia³, and congenital malformations in newborns⁴. Two recent studies^{5,6} have reported, although with some controversy^{7,8}, that germline mutation rates were increased in humans and voles living close to Chernobyl, but little is known about the viability of the organisms affected⁹. Here we report an increased frequency of partial albinism, a morphological aberration associated with a loss of fitness, among barn swallows, *Hirundo rustica*, breeding close to Chernobyl. Heritability estimates indicate that mutations causing albinism were at least partly of germline origin. Furthermore, evidence for an increased germline mutation rate was obtained from segregation analysis at two hypervariable microsatellite loci, indicating that mutation events in barn swallows from Chernobyl were two- to tenfold higher than in birds from control areas in Ukraine and Italy.

The phenotypic consequences of mutation range from altered behaviour and physiology to aberrant morphology¹⁰. Mutations affecting external morphology in birds often lead to partial albinism caused by recessive genes, resulting in partial loss of plumage pigmentation¹¹. Such aberrant coloration is presumed to be detrimental in terms of costs of increased risks for predation and reduced mating success¹¹. Partial albinism in birds appears as spots of white feathers that completely lack pigments. Barn swallows are socially monogamous, aerially insectivorous passerine birds of mass ~20 g, with an annual survival rate of about 28% (ref. 12). Like most other passerines, barn swallows start to breed at the age of one year. European populations migrate across the Sahara desert, wintering in southern Africa. Adult and nestling barn swallows captured around Chernobyl, Ukraine (the radioactively contaminated area), Kanev (an uncontaminated control area in Ukraine), and Milan, Italy (an uncontaminated control area outside the Ukraine) in 1991 and 1996, and museum specimens from before 1986, were inspected for the presence of albinistic feathers in the red front and the blue head, neck, back and tail. There was a significantly higher frequency of partial albinism in Chernobyl than in Kanev swallows in 1991 and 1996 (Fisher exact test; 1991, $P = 0.0000016$; 1996, $P = 0.046$), but not before 1986 ($P = 1.00$; Table 1). Furthermore, there was a significant difference in frequency of partial albinism between swallows from Chernobyl and Italy in 1996 (Fisher exact

test; $P = 0.000772$), but not before 1986 (pre-1986, $P = 1.00$; Table 1). These results indicate that adult barn swallows had an increased frequency of partial albinism in Chernobyl after 1986, but a similar temporal change was not observed in the control area in Ukraine or in Italy.

Partial albinism appears to be associated with a reduced probability of survival. None of 13 partially albinistic adult barn swallows in a Danish study population followed during the years 1985 to 1996 survived from one year to the next¹², whereas 28.0% of 2,110 non-albinistic adults survived (Fisher exact test, $P = 0.0284$). Survival rates based on recaptures are reliable for the barn swallow in this case because almost all surviving adults are captured annually owing to an intensive capture effort, and breeding dispersal of adults is extremely limited¹². A second piece of evidence also suggests a fitness reduction among barn swallows in Chernobyl. The size of the breeding population assessed from the number of nests with eggs or nestlings during the first clutch (a reliable estimate of the size of the breeding population¹²) decreased from 292 pairs in 1991 to 76 pairs in 1996 in the nine villages near Chernobyl (a reduction of 74.0%), but changed rather less, from 202 pairs in 1991 to 162 pairs in 1996, in the six villages in the control area in Ukraine (a reduction of 19.8%). Population size decreased in all nine villages near Chernobyl between 1991 and 1996, whereas only two of six villages near Kanev had a decreasing population (Fisher exact test, $P = 0.022$). Moreover, analyses of developmental instability of feather characters of adult barn swallows around Chernobyl and Kanev in 1991 have similarly indicated that the level of fluctuating asymmetry and the frequency of developmental anomaly was increased considerably in the radioactively contaminated area after 1986 (ref. 13).

Partial albinism in birds can be caused by either somatic or germline mutations¹¹. We compared the parent-offspring resemblance in partial albinism for the sample of birds from Chernobyl in 1991 and 1996. Adult barn swallows develop their plumage in the

African winter quarters, whereas nestlings develop their plumage in the breeding areas. The similarity of the dichotomously classified phenotypes across generations can be used to obtain an estimate of heritability¹⁴. Whereas 12.9% of the offspring with normal-phenotype parents had albinistic feathers ($N = 62$ families), 84.6% of those with partially albinistic parents had such phenotypes ($N = 13$ families). The resemblance was estimated as $r_{(p)} = 0.62$, which differs significantly from zero ($\chi^2 = 25.5$, d.f. = 1, $P < 0.001$; ref. 15). This result indicates that the partially albinistic phenotypes were at least partly determined by germline mutations.

Assessment of the germline mutation rate often proves to be extremely difficult because of the very low frequency of spontaneous mutations seen at most genomic loci. However, two hypervariable (heterozygosity >95%) tetranucleotide microsatellite repeat loci, *HrU6* and *HrU9*, associated with mutation rates of as high as 0.5–3.5% per meiosis¹⁶ (C.R.P., A.P.M. and H.E., unpublished data) have been isolated from the barn swallow genome^{17,18}. We used these markers to genotype family material from Chernobyl, Kanev and Italy (Fig. 1, Table 2). For *HrU6* there was a significantly higher mutation rate in Chernobyl than in Kanev ($\chi^2 = 3.95$, d.f. = 1, $P = 0.047$) and Italy ($\chi^2 = 20.1$, d.f. = 1, $P < 0.001$). The rather small sample sizes from Chernobyl and Kanev make a quantitative estimate of the difference in mutation rate between contaminated and uncontaminated areas only approximate, but it may be as high as an order of magnitude. For *HrU9*, the mutation rate showed a significant two- to threefold increase in Chernobyl compared to Italy ($\chi^2 = 5.00$, d.f. = 1, $P = 0.025$), whereas there was no clear difference in the mutation rate between Chernobyl and Kanev ($\chi^2 = 0.04$, d.f. = 1, $P = 0.91$). Finally, the combined mutation rate for *HrU6* and *HrU9* in Chernobyl was significantly higher than in Italy ($\chi^2 = 16.1$, d.f. = 1, $P < 0.001$) but not Kanev ($\chi^2 = 1.50$, d.f. = 1, $P = 0.22$). Because microsatellite mutation rates may be higher in male than in female germ line¹⁹, and may intrinsically also be dependent on the repeat

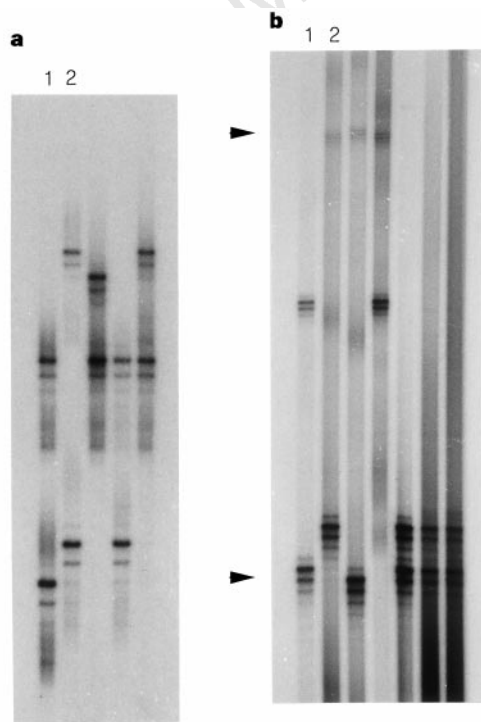


Figure 1 Examples of microsatellite germline mutations for barn swallow loci in the Chernobyl population. **a**, *HrU6*; **b**, *HrU9*. Lane 1, the father; lane 2, the mother; other lanes show offspring. Mutant alleles are arrowed. Note that the offspring to the left in **b** is mutant for both its father's and mother's allele.

Table 1 Frequency of adult barn swallows with partial albinism

Year	Area					
	Chernobyl		Kanev		Milan	
	Albinistic (%)	N	Albinistic (%)	N	Albinistic (%)	N
Pre-1986	0.0	17	0.0	24	0.0	66
1991	15.2	99	0.0	141		
1996	13.3	75	1.9	51	1.7	180

The birds are from Chernobyl, Ukraine (radioactively contaminated area in 1986), Kanev (uncontaminated control area in Ukraine), and Milan, Italy (uncontaminated control area outside Ukraine). Pre-1986 samples of adult barn swallows were obtained from museum collections. Samples from 1991 and 1996 were from mistnet catches of birds at breeding sites.

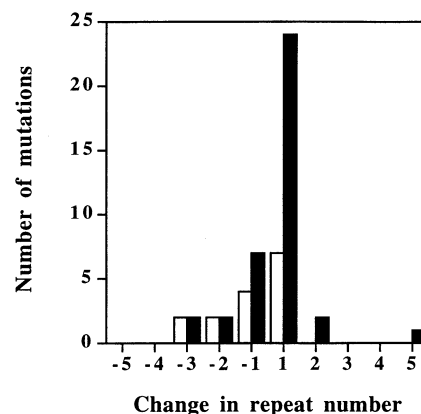


Figure 2 Distribution of changes in the number of repeat units between mutant and progenitor alleles at the barn swallow microsatellite loci *HrU6* and *HrU9*. Populations shown are from Chernobyl (white bars) and Italy (black bars).

Table 2 Frequency of germline mutations at two microsatellite loci

Area	Locus					
	<i>HrU6</i>		<i>HrU9</i>		<i>HrU6 + HrU9</i>	
	Mutations (%)	N	Mutations (%)	N	Mutations (%)	N
Chernobyl	5.6	72	9.1	66	7.2	138
Kanev	0.0	69	8.0	62	3.8	131
Milan	0.5	1,065	3.6	937	2.0	2,002

Adult barn swallows are from Chernobyl, Ukraine (radioactively contaminated area), Kanev (uncontaminated control area in Ukraine) and Milan, Italy (uncontaminated control area outside Ukraine). *N*, number of meioses examined.

length of individual alleles¹⁶, these factors need to match between two populations to allow proper comparisons of mutation rates. However, there was no indication that the increased mutation rate for *HrU6* and *HrU9* in Chernobyl could be explained by population differences in the sex or allele size distributions; the sex ratio (percentage males) among the meioses scored in the Chernobyl group (39.4%) was actually lower than in the Kanev (43.6%; $\chi^2 = 0.48$, d.f. = 1, $P = 0.48$) and Italian (48.3%; $\chi^2 = 4.06$, d.f. = 1, $P = 0.044$) populations. Furthermore, there was no significant difference in the allele frequency distributions between Chernobyl and Kanev for *HrU6* (Kolmogorov–Smirnov, $P = 0.80$) or for *HrU9* ($P = 0.82$), or between Chernobyl and Italy for *HrU6* ($P = 0.41$) or for *HrU9* ($P = 0.54$).

The relatively small sample sizes from Chernobyl and Kanev may explain why we did not observe a significant difference in the mutation rate at *HrU9* between the two areas. However, only two of four human minisatellite systems (two single-locus and two multi-locus probes) had a significantly increased mutation rate in a population from Mogilev, Belarus (a radioactively contaminated area) compared to a control population from United Kingdom⁶. It is not yet clear which other factors may influence microsatellite and minisatellite instability in response to mutagens. Microsatellite mutant alleles typically differ by just one (mainly) or a few repeat units in size from the parental allele^{16,19}, a situation often explained by slipped-strand mispairing during replication²⁰. Mutant alleles from Chernobyl followed a similar distribution of differences in the number of repeat units compared with the parental allele (Fig. 2), and was not significantly different from that observed for mutations in the Italian¹⁶ study population (Kolmogorov–Smirnov, $P = 0.34$). If slippage was also the main mechanism for the observed Chernobyl mutations, it is not likely that this could be associated directly with radiation damage, as ionization generally induces double-strand breaks²¹. Another possibility is damage in repair-systems genes, and in yeast the slippage rate of simple repeats increases dramatically in strains deficient in DNA mismatch repair²². Mutations in mismatch repair genes associated with colorectal cancer and other carcinomas in human and mouse also lead to an increased instability of microsatellites^{23–25}.

Thus both genetic and morphological data suggest that barn swallows breeding in the Chernobyl area have suffered elevated levels of germline mutation rates at both expressed (yielding albinism) and non-coding (microsatellite) loci from the 1986 release of radioactive material from the nuclear power plant. This seems to have resulted in a loss of fitness among individuals in the breeding population, and may also be associated with a significant decline in population size between 1986 and 1996. We have no reason to believe that the barn swallow should be uniquely sensitive to radiation among animal species breeding close to Chernobyl or in other radioactively contaminated areas. □

Methods

Field work. A.P.M. captured barn swallows at farms near Chernobyl (51°17' N, 30°13' E), Ukraine, at a distance of 25–50 km from the nuclear power plant, and around Kanev (49°42' N, 31°25' E), Ukraine, 100 km southeast of Kiev,

outside the contamination zone, on 10–29 June 1991 and 16–24 June 1996. N. Saino captured barn swallows at farms near Milan, Italy, during April–July 1996. The Chernobyl area had an atmospheric level of radiation of 300–500 μ R, which is considerably above the background level (A. A. Tokar, personal communication). Adult barn swallows were captured in mist nets, measured, inspected for partial albinism, ringed and provided with an individual combination of colour on their white breast and belly feathers. A blood sample of 100 μ l was collected from the brachial vein of each individual. Adults were assigned to their nests using observations of birds feeding nestlings, and blood samples of these nestlings were subsequently collected and their plumage inspected for the presence of partial albinism. A.P.M. inspected museum specimens of adult barn swallows from the breeding season in the Natural-History Museums in Kiev, Ukraine and Milan, Florence and Rome, Italy, for the presence of partially albinistic feathers in the plumage. Ukrainian museum samples were divided into birds collected from an area around and northwest of Chernobyl, within the subsequently most heavily contaminated areas, including the area from which field data were sampled in 1991 and 1996, and birds from an area southeast of Kiev (that includes Kanev), well outside the contaminated area¹³. There should be no reason to expect the age distribution among adults to differ between shot birds and breeding birds trapped by us, so the origin and age of pre-1986 Chernobyl birds should match the 1991 and 1996 samples. Moreover, because Ukrainian museum samples consisted of adult birds killed in the field using a shotgun, there is no possibility that the samples were biased with respect to partial albinism. Partial albinism generally involves less than 10 small, white body feathers, and as such is visible only at close range or when the bird is held in the hand.

DNA work. Genomic DNA was prepared by Chelex extraction from about 2 μ l of whole blood. Of the DNA preparations, 0.1–0.5% were used as template in 10 μ l PCR reactions with *HrU6* or *HrU9* containing 1 pmol of each primer (one end-labelled with [γ -³²P]ATP), 1.5 mM MgCl₂, 50 mM KCl, 10 mM Tris, 200 μ M dNTP and 0.5 U *Taq* polymerase (Dynazyme). After initial denaturation at 90 °C for 3 min, 30 cycles of 94 °C for 30 s, 63 °C for 30 s and 72 °C for 80 s were run. PCR products were separated in 6% polyacrylamide gels and visualized by autoradiography. Known size standards were included at regular intervals on the gels. Because the allele size range of *HrU6* and *HrU9* spans several hundred base pairs, extended electrophoresis was performed to achieve appropriate separation of long alleles. Identification of mismatching alleles was facilitated by the fact that *HrU6* and *HrU9* are both tetranucleotide repeats, and so most alleles differ by at least four base pairs in size. An offspring allele not compatible with the alleles of either of its presumed parents could indicate either a mutation event (relevant both for maternal and paternal transmission) or false parentage (extra-pair paternity, EPP). Distinguishing between these two events is critical to this study because EPP is not uncommon in barn swallows²⁶. To identify true genetical relationships accurately, an extended paternity analysis was therefore conducted by genotyping a set of eight additional, polymorphic microsatellites *HrU2-5*, *HrU7-8* (ref. 17), *HrU10* (ref. 18) and *Escμ6* (ref. 27). In 10 cases for which one or two alleles at either *HrU6* or *HrU9* did not match with the presumed mother and/or father, but that of the other locus did, the segregation at all additional eight markers was compatible with a correct maternity and paternity assignment, strongly indicating that the mismatch at *HrU6* or *HrU9* was due to mutation rather than EPP. According to observed allele frequency distributions, the mean probability that a random individual would match at least 9 of our 10 loci is as low as 0.0005 (95% confidence interval of 0.001–0.0001). In the sample of 1996 Ukrainian offspring, we would have expected to find only 0.04 birds where an allelic mismatch at one of the 10 loci would be due to EPP rather than mutation. We consider this possible effect negligible, and conclude that 10 microsatellite mutations were present in the sample. For 24 offspring more than one marker did not match with the presumed father (one case with four markers excluding the associated male, five cases with five markers, nine cases with six markers, four cases with seven markers and five cases with eight markers), and these offspring were inferred as instances of EPP. The frequency of such instances in the Ukrainian material (30%) was similar to that in our study populations in Italy²⁶ and Denmark¹⁷.

Received 3 June; accepted 4 August 1997.

1. Ansbaugh, L. R., Catlin, R. J. & Goldman, M. The global impact of the Chernobyl reactor accident. *Science* **242**, 1513–1519 (1988).

2. Kazakov, V. S., Demidchik, E. P. & Astakhova, L. N. Thyroid cancer after Chernobyl. *Nature* **359**, 21 (1992).
3. Petridou, E. *et al.* Infact leukaemia after in utero exposure to radiation from Chernobyl. *Nature* **382**, 352–353 (1996).
4. Lazjuk, G. I., Kirillova, I. A., Dubrova, Y. E. & Novikova, I. V. in *The Chernobyl Papers Vol. 1, Doses to the Soviet Population and Early Health Effects Studies* (eds Mervin, S. E. & Balanov, M. I.) 385–397 (Research Enterprises, Richland, WA, 1993).
5. Baker, R. J. *et al.* High levels of genetic change in rodents of Chernobyl. *Nature* **380**, 707–708 (1996).
6. Dubrova, Y. E. *et al.* Human minisatellite mutation rate after the Chernobyl accident. *Nature* **380**, 683–686 (1996).
7. Dubrova, Y. E. *et al.* Effects of radiation on children (reply). *Nature* **383**, 226 (1996).
8. Satoh, C. & Kodaira, M. Effects of radiation on children. *Nature* **383**, 226 (1996).
9. Hillis, D. M. Life in the hot zone around Chernobyl. *Nature* **380**, 665–666 (1996).
10. Auerbach, C. *Mutation Research* (Chapman & Hall, London, 1976).
11. Dyck, J. in *A Dictionary of Birds* (eds Campbell, B. & Lack, E.) 472–474 (Poyser, Calton, 1985).
12. Möller, A. P. *Sexual Selection and the Barn Swallow* (Oxford Univ. Press, 1994).
13. Möller, A. P. Morphology and sexual selection in the barn swallow *Hirundo rustica* in Chernobyl, Ukraine. *Proc. R. Soc. Lond. B* **252**, 51–57 (1993).
14. Falconer, D. S. *Introduction to Quantitative Genetics* 2nd edn (Longman, New York, 1989).
15. Siegel, S. & Castellan, N. J. Jr *Nonparametric Statistics for the Behavioral Sciences* 2nd edn (McGraw-Hill, New York, 1988).
16. Primmer, C. R., Ellegren, H., Saino, N. & Möller, A. P. Directional evolution in germline microsatellite mutations. *Nature Genet.* **13**, 391–393 (1996).
17. Primmer, C. R., Möller, A. P. & Ellegren, H. Resolving genetic relationships with microsatellite markers: a parentage testing system for the swallow *Hirundo rustica*. *Mol. Ecol.* **4**, 493–498 (1995).
18. Primmer, C. R., Möller, A. P. & Ellegren, H. New microsatellite markers from the pied flycatcher *Ficedula hypoleuca* and barn swallow *Hirundo rustica* genomes. *Heredity* **124**, 281–283 (1996).
19. Weber, J. L. & Wong, C. Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**, 1123–1128 (1993).
20. Levinson, G. & Gutman, G. A. Slipped-strand mis-pairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203–221 (1987).
21. Ward, J. F. The yield of DNA double-strand breaks produced intracellularly by ionizing radiation: a review. *Int. J. Radiat. Biol.* **57**, 1141–1150 (1990).
22. Strand, M., Prolla, T. A., Liskay, R. M. & Petes, T. D. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **365**, 274–276 (1993).
23. Parsons, R. *et al.* Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell* **75**, 1227–1236 (1993).
24. Fishel, R. *et al.* Binding of mismatched microsatellite DNA sequences by the human MSH2 protein. *Science* **266**, 1403–1405 (1994).
25. Reitmaier, A. H. *et al.* MSH2 deficient mice are viable and susceptible to lymphoid tumours. *Nature Genet.* **11**, 64–70 (1995).
26. Saino, N., Primmer, C. R., Ellegren, H. & Möller, A. P. An experimental study of paternity and tail ornamentation in the barn swallow (*Hirundo rustica*). *Evolution* **51**, 562–570 (1997).
27. Hanotte, O. *et al.* Isolation and characterization of microsatellite loci in a passerine bird: the reed bunting *Emberiza schoeniclus*. *Mol. Ecol.* **3**, 529–530 (1994).

Acknowledgements. We thank A. A. Tokar for logistic support in Ukraine; the curators of museum collections in Kiev, Ukraine, and Milan, Florence and Rome, Italy for access to specimens; and N. Saino for providing phenotypic data from his barn swallow population in 1996. A.P.M. was supported by the Danish Natural Science Research Council, H.E. by the Swedish Research Councils for Natural Sciences, and for Agriculture and Forestry and G.L. by the Elis Wides Fund (Swedish Ornithological Society).

Correspondence and requests for materials should be addressed to H.E. or A.P.M. (respective e-mail addresses: hans.ellegren@bmc.uu.se or anders.moller@snv.jussieu.fr).

How the brain learns to see objects and faces in an impoverished context

R. J. Dolan^{††}, G. R. Fink^{*}, E. Rolls[‡], M. Booth[‡], A. Holmes^{*}, R. S. J. Frackowiak^{*} & K. J. Friston^{*}

^{*} Wellcome Department of Cognitive Neurology, Institute of Neurology, Queen Square, London W1N 3BG, UK

[†] Academic Department of Psychiatry, Royal Free Hospital School of Medicine, London NW3, UK

[‡] Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

A degraded image of an object or face, which appears meaningless when seen for the first time, is easily recognizable after viewing an undegraded version of the same image¹. The neural mechanisms by which this form of rapid perceptual learning facilitates perception are not well understood. Psychological theory suggests the involvement of systems for processing stimulus attributes, spatial attention and feature binding², as well as those involved in visual imagery³. Here we investigate where and how this rapid perceptual learning is expressed in the human brain by using functional neuroimaging to measure brain activity during exposure to

degraded images before and after exposure to the corresponding undegraded versions (Fig. 1). Perceptual learning of faces or objects enhanced the activity of inferior temporal regions known to be involved in face and object recognition respectively^{4–6}. In addition, both face and object learning led to increased activity in medial and lateral parietal regions that have been implicated in attention⁷ and visual imagery⁸. We observed a strong coupling between the temporal face area and the medial parietal cortex when, and only when, faces were perceived. This suggests that perceptual learning involves direct interactions between areas involved in face recognition and those involved in spatial attention, feature binding and memory recall.

The experimental design involved four conditions: two-tone (black and white) images of objects before (Ob) and after (Oa) exposure to the associated grey-scale images of the objects; two-tone images of faces before (Fb) and after (Fa) exposure to the associated grey-scale images of the faces. A factorial experimental design, crossing pre- and post-grey-scale exposure with object or face perception, ensured that the same visual input was maintained across all levels of stimulus presentation, with the condition-specific neural responses being measured with positron emission tomography (PET) indices of local perfusion. Thus, the factors in the experiment are exposure (two levels: before and after learning) and stimulus (two levels: object and face). Each combination of levels was repeated 3 times on 8 subjects (study 1; $n = 96$) and 6 subjects (study 2; $n = 72$) as described below.

In study 1, we first examined the main effect of exposure using the contrast (Oa + Fa) – (Ob + Fb) to identify regions activated by the combined detection of objects and faces in study 1. This comparison revealed extensive bilateral activation in the medial parietal region ($P < 0.05$, corrected), corresponding to the precuneus, and extending laterally to involve the posterior inferior parietal cortex ($P < 0.05$, corrected; Fig. 2).

We next examined the main effect of stimulus to localize category-specific activations. Comparing object with face conditions, using the contrast (Ob + Oa) – (Fb + Fa), showed activations in the left inferior temporal region ($P < 0.05$, corrected) (Fig. 3a). The opposite contrast (Fb + Fa) – (Ob + Oa) revealed activation in the right superior temporal region ($P < 0.05$, corrected) extending into the right inferior temporal region ($P < 0.001$, uncorrected) in association with the face conditions (Fig. 3b). Finally, to identify regions whose object-specific responses were enhanced by exposure-dependent perception, we tested for the conjunction of object-specific responses (that is, the main effect of objects relative to faces, which equals (Ob + Oa) – (Fb + Fa)) and significantly increased perception-dependent activations (that is, category $\times$ exposure interaction = (Ob + Oa) – (Fb + Fa)). In this analysis, learning-dependent effects due to object perception were seen in the left fusiform region ($z = 4.4$; $P < 0.05$ corrected) (Fig. 3c). The equivalent conjunction analysis for learning-dependent effects due to faces demonstrated a similar effect in the right fusiform gyrus (Fig. 3d) ($z = 3.22$; $P < 0.001$, uncorrected).

The experimental design interposes an explicit learning phase between first and second stimulus presentation and includes a learning component due to repeated exposure to the same stimuli. To remove the influence of this order-effect learning, we contrasted the patterns of activation associated with explicit learning (study 1) with those elicited in a second experiment (study 2) by interposing non-associated grey-scale images. These contrasts involved a group $\times$ effects interaction where the effects comprised those associated with category-independent learning (combined post-exposure minus pre-exposure) and those specific to learning in the faces and objects conditions (post-exposure minus pre-exposure for faces minus objects and vice versa). The comparisons reach significance at an uncorrected level and are reported descriptively. The group $\times$ exposure interaction (category-independent learning)

identified the medial parietal cortex ($z = 2.9$; $P = 0.002$, uncorrected). For the *group* $\times$ *exposure* $\times$ *stimulus* interaction (category-specific learning) the right fusiform gyrus activation was significant ($z = 3.6$; $P < 0.001$, uncorrected) for face learning but the left fusiform activation for object learning failed to reach significance even at a threshold $z = 2.33$ (corresponding to $P < 0.01$).

We next tested the hypothesis that perceptual learning involved interactions between the medial parietal cortex, an area implicated in higher-order cognition⁸, and the area showing face-learning responses. We tested this hypothesis by creating a statistical parametric map to identify areas where activity can be explained in terms of face-specific interactions with medial parietal cortex activity (Fig. 2). The analysis used a statistical model that includes an effect of faces relative to objects (and vice versa) and a term that represents the interaction between this effect and adjusted activity in the medial parietal cortex. This term is simply the adjusted activity in the parietal region multiplied by 1 when faces were present and -1 when they were not. The significance of this interaction term was assessed using *t*-statistics and an uncorrected threshold of $P < 0.01$. Because we predicted that this modulatory effect would be expressed in the fusiform gyrus, we selected the regional activation (surviving a threshold of $P < 0.01$, uncorrected) that was closest to the boundary of the fusiform gyrus. This region comprised 340 voxels. The probability of this number of voxels or more being obtained by chance was $P < 0.039$; $z = 4.0$. This can be considered as a *P* value that has been corrected for the volume of the fusiform gyrus, indicating a significant modulation of face-specific responses in the right fusiform region by medial parietal cortex⁹ (Fig. 4).

We have made several critical findings. First, learning-facilitated

perception of objects or faces in degraded stimuli specifically involves activation of medial parietal cortex. Second, we observed enhanced neural responses in the fusiform gyri to an identical visual input when it elicited a strong visual percept. This effect was most striking for the enhanced percept of faces, for which augmentation of activation in the right fusiform involved a significant interaction with medial parietal cortex. The lack of significance for a similar effect in the object condition, in the between-group analysis, we attribute to greater nonspecific effects of attentional set engendered when subjects attempt a reconstruction of objects following exposure to non-related grey-scale objects in the second study.

Category-dependent lateralized activations, in the direct comparisons of object and face conditions, do not necessarily imply that object and face processing solely involve the left and right fusiform regions. We would expect bilateral activations if we had included a neutral control condition. Selective fusiform activations during face perception have been reported^{10–12}. Likewise, inferior temporal cortex and nearby occipital regions play a role in pattern and object recognition^{13–17}. Our findings extend these observations by demonstrating that the right fusiform region alters its response to activation as a result of learning. Moreover, as the same visual inputs were used across all conditions, our findings specifically address what brain areas are activated when faces and objects are perceived.

Another novel finding is the observation that learning-facilitated perception involves not only fusiform regions but also medial (specific to learning) and lateral parietal cortex. The functional role of the latter regions in learning-facilitated perception is open to interpretation. First, perception-related fusiform activations may reflect learning-related tuning of neural activity similar to that seen in monkeys^{18–22}. The parietal responses could reflect propagated

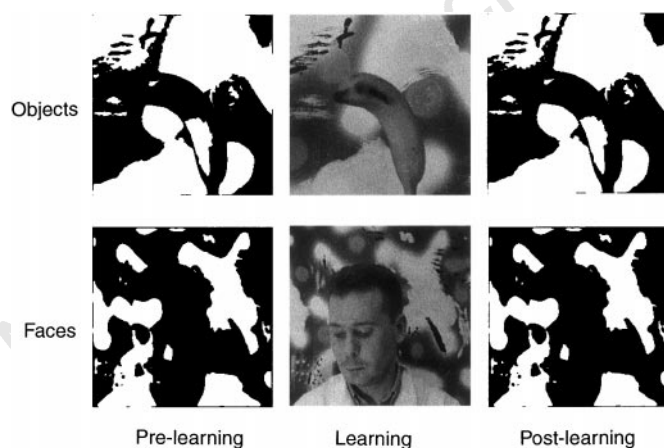


Figure 1 The experimental design. Binarized images of an object (top row) and face (bottom row) and their associated full grey-scale versions. The process of binarizing images involved transforming grey-scale levels into either black or white (two-tone) with values of either 0 or 1. Exposure to the associated grey-scale version took place 5 min before a second exposure to the two-tone version in study 1. In the pre- and post-learning scans, the subjects were told that they might see faces or objects, respectively, in the stimuli. A significant behavioural learning effect, operationally defined as the facilitation of performance by prior exposure to the grey-scale version, was evident for the object and face conditions. Behavioural data were collected for all conditions at the end of each individual scan. The mean number of resolved percepts pre- and post-exposure to the grey-scale images were 13 and 87% in the object, and 55 and 93% in face conditions, respectively. In study 2, the same sequence and stimuli were used but with pre- and post-exposure being interposed with non-related grey-scale images of objects and faces. Here the mean number of resolved percepts pre- and post-exposure to non-associated grey-scale images were 8 and 10% in the object, and 19 and 25% in the face conditions, respectively.

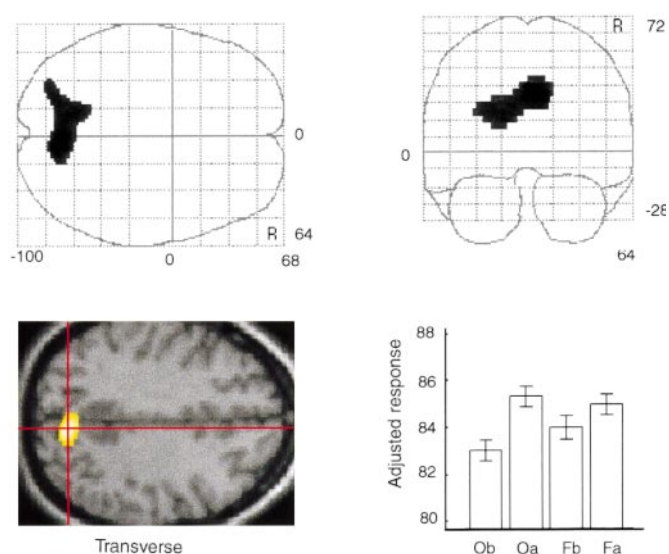


Figure 2 Main effect of exposure. Relative activations (for the eight subjects) associated with post-learning perception of objects or faces; that is, main effect of exposure ($Oa + Fa - (Ob + Fb)$). Areas of significant activations ($P < 0.05$, corrected) are shown as through-projection onto representations of standard stereotaxic space. Coronal, back view; transverse, view from above. Transverse SPM(*z*) maps were superimposed upon the group mean magnetic resonance image (MRI) spatially normalized into the same anatomical space. Coordinates are described with reference to the Talairach system³⁰. The local maximum within the activated volume is defined by the intersection of red lines. There is bilateral activation centred on the precuneus ($x = 4$ mm, $y = -74$ mm, $z = 36$ mm; $P < 0.05$, corrected; z -score = 5.7), extending into the lateral aspects of the inferior posterior parietal cortex. Adjusted mean rCBF (to a mean of $50 \text{ ml dl}^{-1} \text{ min}^{-1}$) with s.e.m. per condition displayed for the local maximum at the above coordinates.

outputs from fusiform regions that engage higher-order visual regions to mediate a spatial reconstruction, from fragmentary evidence, of perceived images. The lateral parietal cortex is implicated in spatial attention and in feature binding necessary for spatial representation of objects²³. The medial parietal cortex (precuneus) on the other hand, is implicated in memory-related imagery^{8,24}. As our experimental tasks involve reconstruction of object (or face) representations from fragmentary evidence we suggest that exposure-related activation of medial parietal cortex reflects this 'mind's eye' reconstruction which may be a necessary process in high-level perception.

An alternative account is that category-specific enhanced fusiform activations reflect neuronal populations modulated by top-down influences, from the parietal cortex, representing an interaction of mnemonic, imagery and attentional processes with category-specific stimuli. This synthesis would necessitate on-going

dynamic interactions between category-specific regions (fusiform gyri) and non-category-specific higher-order regions (parietal) necessary for the coherent synthesis of a percept from an insufficiently specified input. Top-down effects are implicated by psychophysical experiments that demonstrate imagery-based facilitation of early sensory representations²⁵. Likewise, single-unit recordings demonstrate that memory and attention can tune responses of inferior temporal cortex to repeated encounters with stimuli⁴⁻⁶, and neuroimaging data has shown that directed attention influences early sensory processing^{26,27}. The likely mechanisms can be inferred from experiments that demonstrate behaviourally specific reciprocal interactions between parietal and visual regions, involving synchronization of responses, in the awake cat²⁸.

The striking modulatory influences demonstrated here indicate that learning-based facilitation of perception involves response modulation in category-specific extrastriate neuronal populations

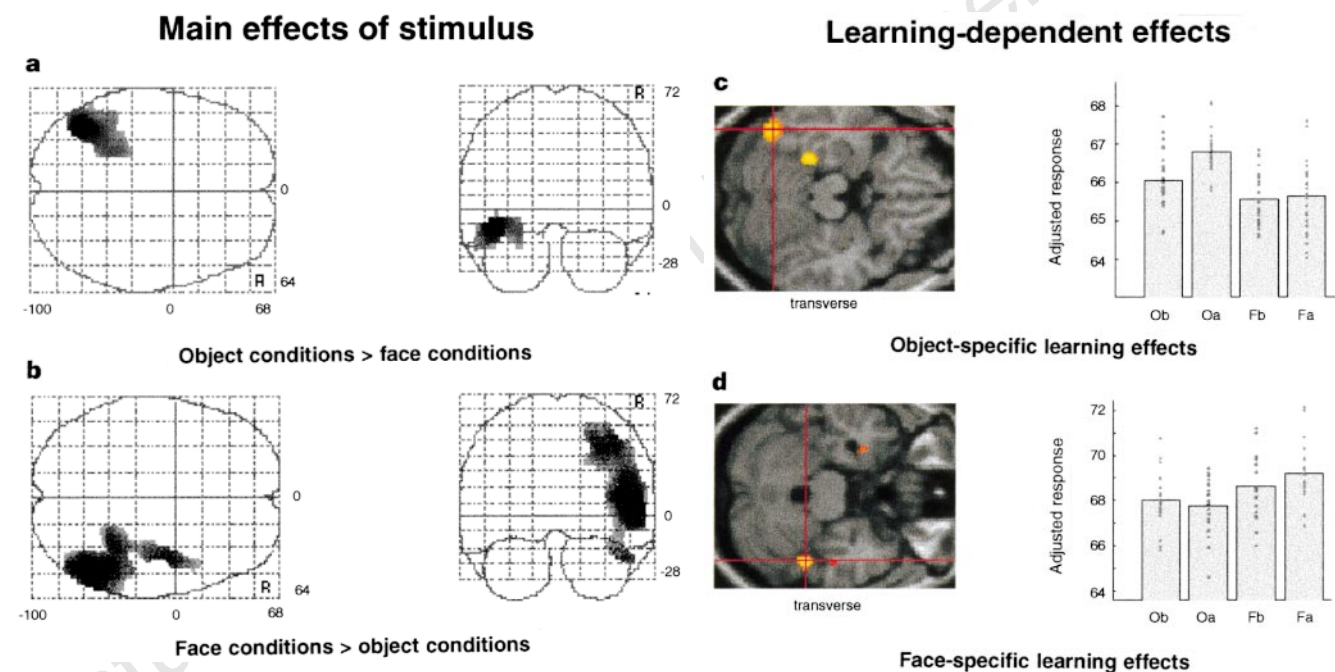


Figure 3 Main effects of stimulus in study 1 for **a**, objects (Ob + Oa) - (Fb + Fa), and **b**, faces (Fb + Fa) - (Ob + Oa) according to the convention of Fig. 2 and in the text. **c, d**, Regions showing category-specific learning effects for objects and faces as described in the text, superimposed on to transverse MRI sections. Local maxima are defined by the intersection of the red lines corresponding to

fusiform gyrus on the left for object (**c**) ($x = -46$ mm, $y = -60$ mm, $z = -16$ mm; z -score = 4.4, $P < 0.05$, corrected) and right for faces (**d**) ($x = 44$ mm, $y = -38$ mm, $z = -28$ mm; z -score = 3.2, $P < 0.001$). Adjusted mean rCBF (mean of $50 \text{ ml dl}^{-1} \text{ min}^{-1}$), with individual data points per condition for these local maxima centred on the above co-ordinates.

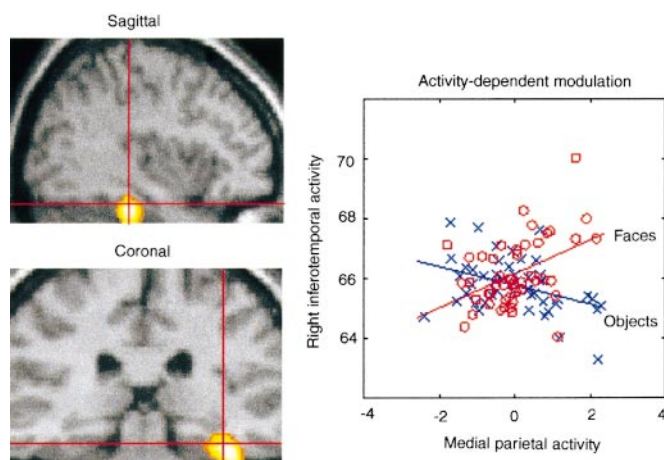


Figure 4 Left, regions in the right fusiform gyrus where there are significant modulatory effects from the medial parietal cortex; right, same effect as individual data points with respect to the face and object condition. Circles and crosses represent activity during the presentation of faces and objects, respectively. rCBF equivalents for the right fusiform gyrus are plotted as a function of activity in the specified medial parietal region ($x = 4$ mm, $y = -74$ mm, $z = 36$ mm). It can be seen that the right fusiform region responds to faces (relative to objects) when, and only when, parietal activity is high. This analysis directly confirms that face-specific activity can be sufficiently explained by category-specific responses that are modulated by parietal influences extant at the time of stimulus presentation.

and interactions with higher-order brain regions. Our results support psychological theories that perception is a conjoint function of current sensory input interacting with memory and possibly attentional processes^{2,3}. The findings extend these psychological observations by providing neurobiological evidence that memory and attention organize the visual input through selective effects on extrastriate visual processing. □

Methods

In study 1, eight normal-sighted male volunteers (six right- and two left-handed), and six normal-sighted right-handed male volunteers in study 2 were scanned using a SIEMENS/CPS ECAT EXACT HR+ (model 962) PET scanner (Siemens/CTI) in 3-dimensional mode with a 15-cm axial field of view. Relative cerebral blood flow (rCBF) was measured from the distribution of radioactivity after slow bolus intravenous (i.v.) injection of H₂¹⁵O (9 mCi per scan, each lasting 90 s). Attenuation-corrected data were reconstructed into 63 image planes with a resulting resolution of 6 mm at full-width-half-maximum. For each subject, structural magnetic resonance (MR) images were obtained with a 2 T Magnetom VISION scanner (Siemens).

Twelve regional perfusion measurements were taken during the following four conditions: two-tone images of objects before learning (the learning phase consisted of presentation of the grey-scale images); two-tone images of objects post-learning; two-tone images of faces pre-learning; two-tone images of faces post-learning. The order of presentation of stimuli that represented either objects or faces was counterbalanced but each set of binarized images were grouped back to back. For any set of binarized stimuli, subjects were scanned twice (pre- and post-exposure to the associated grey-scale image). Subjects were aware that binarized images following exposure to faces or objects might relate to the category of stimulus seen in the learning phase. Subjects lay with eyes open in a quiet, darkened room watching a 36-cm video display unit at a viewing distance of 40 cm. Subjects were exposed to individual images for 3 s. Ten images in all were viewed during a single scan window, with a 1-s interstimulus interval. Subjects had 6 scans before and 6 scans after learning (3 scans for objects and 3 for faces, pre- and post-learning). Learning with individual grey-scale images occurred 5 min before re-exposure to the related two-tone images.

Statistical parametric mapping (SPM96, Wellcome Department of Cognitive Neurology, London) software was used for image realignment, transformation into standard stereotactic space, smoothing and statistical analysis²⁹. All measurements per condition were averaged across subjects. State-dependent differences in global flow were co-varied out using ANCOVA. Main effects and interactions were assessed with contrasts of the adjusted task means using *t*-statistics subsequently transformed into the *z* statistic. For between-group comparisons, although our block ANCOVA removed random effects due to subjects, this does not constitute a random effects model. Accordingly, we ensured the appropriateness of our fixed effects model by testing explicitly for *subject* × *contrast* interaction effects and demonstrating that they were not significant. Corrected values refer to correction for the whole brain volume based upon magnitude of the response. The resulting set of *z* values constituted a statistical parametric map (SPM{*z*}). Localization of maxima are reported within the standard space, as defined by Talairach and Tournoux, and were superimposed on the group mean magnetic resonance image spatially normalized into the same anatomical space³⁰.

Received 20 March; accepted 22 July 1997.

1. Ramachandran, V. S. in *The Artful Eye* (eds Gregory, R. L. & Harris, J.) 249–267 (Oxford University Press, Oxford, 1994).
2. Treisman, A. Features and objects. *Quart. J. Exp. Psych.* **40** A, 201–237 (1988).
3. Kosslyn, S. M. *Image and Brain: The Resolution of the Imagery Debate* (MIT Press, Cambridge, MA, 1996).
4. Moran, J. & Desimone, R. Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784 (1985).
5. Baylis, G. C. & Rolls, E. T. Responses of neurones in the inferior temporal cortex in short term and serial recognition memory tasks. *Exp. Brain Res.* **65**, 614–622 (1987).
6. Gross, C. G., Bender, D. B. & Gerstein, G. L. Activity of inferior temporal neurons in behaving monkey. *Neuropsychologia* **17**, 215–229 (1979).
7. Posner, M. I. & Petersen, S. E. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**, 25–42 (1990).
8. Fletcher, P. et al. The mind's eye—activation of the precuneus in memory related imagery. *Neuroimage* **2**, 196–200 (1995).
9. Friston, K. J. Testing for anatomically specified regional effects. *Human Brain Mapping* **5**, 133–166 (1997).

10. Haxby, J. et al. The functional organisation of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.* **14**, 6336–6353 (1994).
11. Clark, V. P. et al. Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. *Neuroimage* **4**, 1–15 (1996).
12. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
13. Desimone, R. & Gross, C. G. Visual areas in the temporal lobe of the macaque. *Brain Res.* **178**, 363–380 (1979).
14. Tanaka, K., Saito, H., Fukada, Y. & Morioka, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurosci.* **6**, 134–144 (1991).
15. Gross, C. G. in *Handbook of Sensory Physiology* (ed. Jung, B. R.) Vol. 7, b; 451–482 (Springer, Berlin, 1972).
16. Ungerleider, L. G. & Mishkin, M. in *Analysis of Behaviour* (eds Goodale, M. A. & Mansfield, R. J. Q.) 549–586 (MIT Press, Cambridge, MA, 1982).
17. Malach, R. et al. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl Acad. Sci. USA* **92**, 8135–8139 (1995).
18. Sakai, K. & Miyashita, Y. Neural organization for the long-term memory of paired associates. *Nature* **354**, 152–155 (1995).
19. Tovee, M. J., Rolls, E. T. & Ramachandran, V. S. Visual learning in neurons of the primate temporal visual cortex. *Neuroreport* **7**, 2757–2760 (1996).
20. Sakia, K. & Miyashita, Y. Neuronal tuning to learned complex forms in vision. *Neuroreport* **5**, 829–832 (1994).
21. Miller, E. K., Li, L. & Desimone, R. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254**, 1377–1379 (1991).
22. Li, L., Miller, E. K. & Desimone, R. The representation of stimulus familiarity in the anterior inferior temporal cortex. *J. Neurophysiol.* **69**, 1918–1929 (1993).
23. Friedman-Hill, S. R., Robertson, L. C. & Treisman, A. Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science* **269**, 853–855 (1995).
24. Shallice, T. et al. Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* **368**, 633–635 (1994).
25. Ishai, A. & Sagi, D. Common mechanisms of visual imagery and perception. *Science* **268**, 1772–1774 (1995).
26. Heinze, H. J. et al. Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* **372**, 543–546 (1994).
27. Fink, G. R. et al. Where in the brain does visual attention select the forest and trees? *Nature* **382**, 626–629 (1996).
28. Roelfsema, P. R., Engel, A. K., Konig, P. & Singer, W. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature* **385**, 157–161 (1997).
29. Friston, K. et al. Statistical parametric mapping in functional imaging: a general linear approach. *Human Brain Mapping* **2**, 189–210 (1995).
30. Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain* (Thieme, Stuttgart, 1988).

Acknowledgements. We thank our volunteers and the radiography staff at the Wellcome Department. R.J.D., K.J.F., R.S.J.F. and G.R.F. are supported by the Wellcome Trust. The bibliographic support of R. Lai continues to be invaluable.

Correspondence and requests for materials should be addressed to R.J.D. (e-mail: r.dolan@fil.ion.ucl.ac.uk).

A hippocampal GluR5 kainate receptor regulating inhibitory synaptic transmission

Vernon R. J. Clarke[†], Barbara A. Ballyk[†], Ken H. Hoo[†], Allan Mandelzys[†], Andrew Pellizzari[†], Catherine P. Bath[‡], Justyn Thomas[‡], Erica F. Sharpe[‡], Ceri H. Davies[§], Paul L. Ornstein^{||}, Darryle D. Schoepp^{||}, Rajender K. Kamboj[†], Graham L. Collingridge^{*}, David Lodge[‡] & David Bleakman[‡]

^{*} Department of Anatomy, University of Bristol, University Walk, Bristol BS8 1TD, UK

[†] Allelix Biopharmaceuticals, 6850 Goreway Drive, Mississauga, Ontario, L4V 1V7 Canada

[‡] Eli Lilly and Company Ltd, Lilly Research Centre, Erl Wood Manor, Windlesham, Surrey GU20 6PH, UK

[§] Department of Pharmacology, University of Edinburgh, 1 George Square, Edinburgh EH8 9JZ, UK

^{||} Eli Lilly and Company Ltd, Lilly Corporate Centre, Indianapolis, Indiana 46285, USA

The principal excitatory neurotransmitter in the vertebrate central nervous system, L-glutamate, acts on three classes of ionotropic glutamate receptors, named after the agonists AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole-4-propionic acid), NMDA (N-methyl-D-aspartate) and kainate¹. The development of selective pharmacological agents has led to a detailed understanding of the physiological and pathological roles of AMPA and NMDA

receptors²⁻⁸. In contrast, the lack of selective kainate receptor ligands has greatly hindered progress in understanding the roles of kainate receptors^{9,10}. Here we describe the effects of a potent and selective agonist, ATPA ((*RS*)-2-amino-3-(3-hydroxy-5-*tert*-butylisoxazol-4-yl)propanoic acid) and a selective antagonist, LY294486 ((3*SR*, 4*aRS*, 6*SR*, 8*aRS*)-6-(((1*H*-tetrazol-5-yl)methyl)oxy)methyl)-1, 2, 3, 4, 4*a*, 5, 6, 7, 8, 8*a*-decahydroisoquinoline-3-carboxylic acid), of the GluR5 subtype of kainate receptor¹¹. We have used these agents to show that kainate receptors, comprised of or containing GluR5 subunits, regulate synaptic inhibition in the hippocampus, an action that could contribute to the epileptogenic effects of kainate¹²⁻¹⁷.

ATPA (Fig. 1a) was developed several years ago as an AMPA receptor agonist, having greater brain penetration than AMPA, despite being intrinsically less potent than AMPA itself^{18,19}. However, its activity at the subtypes of non-NMDA receptors was unknown. We have performed ligand-binding and electrophysiological studies across a range of cloned human AMPA and kainate receptor subtypes stably expressed in HEK293 cells (Fig. 1b). In binding studies, ATPA had weak activity (K_i values of 6–14 μ M) at AMPA receptors (GluR1–4) and the kainate receptors KA-2 and GluR7, but was devoid of activity at GluR6 kainate receptors (>1 mM). However, at GluR5 kainate receptors, ATPA bound with low nanomolar potency (4.3 ± 1.1 nM, mean $\pm$ s.e.m., $n = 4$), making it more potent than kainate itself. In whole-cell recordings from HEK293 cells expressing GluR5 receptors²⁰, ATPA evoked inward currents (EC_{50} of 2.1 ± 0.1 μ M, $n = 6$), as did kainate (EC_{50} of 16.2 ± 1.0 μ M, $n = 6$; Fig. 2a). In contrast, ATPA only weakly activated GluR1–4 receptors (for example, EC_{50} for GluR1 of 386 ± 57 μ M, $n = 4$, and for GluR4 of 662 ± 165 μ M, $n = 4$), and had no effect on GluR6 receptors (100 μ M to 10 mM, $n = 4$).

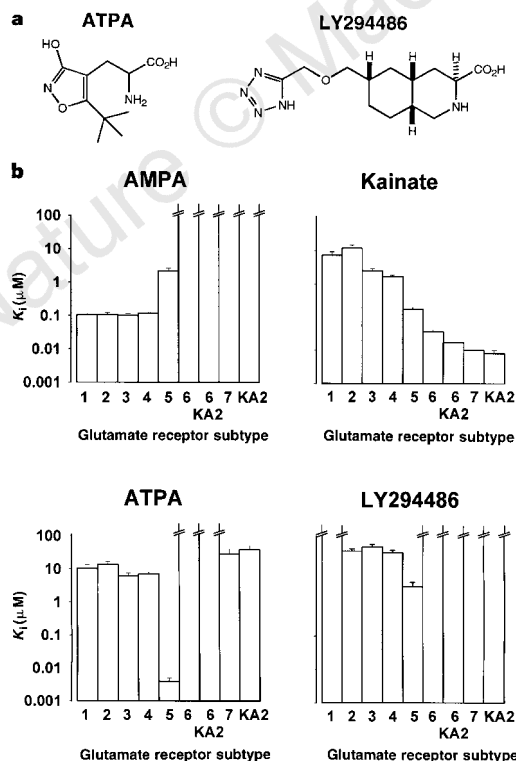


Figure 1 **a**, Structures of ATPA and LY294486. **b**, Pharmacological selectivity of AMPA, kainate, ATPA and LY294486, for AMPA and kainate receptor subtypes using radioligand binding. K_i values were determined from 11 point competition assays. Values shown represent mean K_i values from 3–6 separate preparations. Open tops to the histograms indicate binding affinity >100 μ M.

These results identify ATPA as a highly selective and potent agonist at cloned human GluR5.

Two preparations were used to compare the functional activity of ATPA at native rat neuronal kainate and AMPA receptors with that at cloned human receptors (Fig. 2). Currents activated by kainate in acutely isolated dorsal root ganglion (DRG) neurons are thought to be mediated by GluR5 receptors^{21,22}, whereas currents activated by kainate in acutely isolated cerebellar Purkinje neurons are mediated by AMPA receptors²³. In whole-cell recordings from DRG neurons, ATPA evoked inward currents (EC_{50} of 0.6 ± 0.1 μ M, $n = 4$) with a higher potency than kainate (EC_{50} 12 ± 2 μ M, $n = 4$; Fig. 2b). In whole-cell recordings from cerebellar Purkinje neurons, ATPA, kainate and AMPA all evoked inward currents (EC_{50} values of 340 ± 26 μ M, $n = 3$; 64 ± 10 μ M, $n = 4$ and 2.8 ± 0.4 μ M, $n = 3$, respectively; Fig. 2c). Thus the GluR5 receptor potency and selectivity demonstrated by ATPA at cloned human glutamate receptors extends to native receptors on rat neurons.

We have described a compound, LY293558 (ref. 23), that inhibits GluR5 but not GluR6. However, this compound is also an AMPA receptor antagonist. Other members of this family of decahydroisoquinolines were therefore examined, and a related compound, LY294486 (Fig. 1a), was identified as having greater selectivity for GluR5 over AMPA receptors than LY293558, but still lacking activity at GluR6, GluR7 and KA-2 receptors (Fig. 1b). Good agreement between binding data and functional studies also extends to LY294486 (Fig. 3). Thus LY294486 inhibited kainate responses in cells expressing human GluR5 with a low micromolar IC_{50} (3.9 ± 0.5 μ M, $n = 4$) and in cells expressing GluR1–4 with IC_{50} s greater than 30 μ M. At 100 μ M, LY294486 had no effect on responses evoked from cells expressing GluR6. An example of this potent inhibition of GluR5 is shown in Fig. 3. The receptor profile of activity for LY294486 at recombinant human glutamate receptors was paralleled in native rat neuronal preparations. LY294486 potently inhibited kainate (30 μ M)-evoked responses from DRG neurons (IC_{50} of 0.62 ± 0.14 μ M, $n = 5$; Fig. 3) and ATPA (3 μ M)-evoked responses in these cells (IC_{50} of 1.3 ± 0.2 μ M, $n = 5$). Inhibition of AMPA-evoked responses in Purkinje cells was only

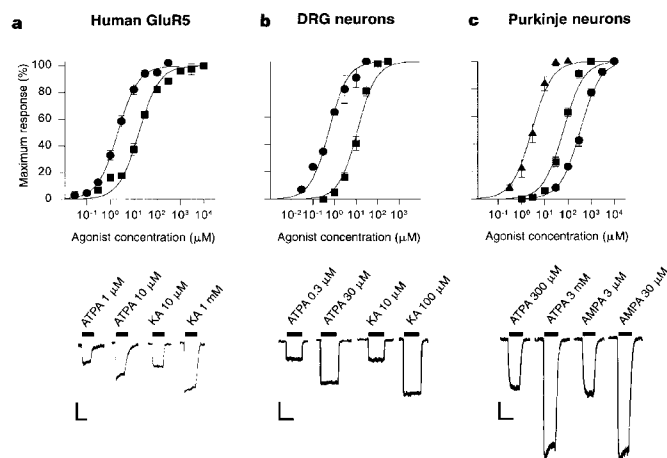


Figure 2 **a**, Agonist concentration-response curves of ATPA (circles) and kainate (squares) at human GluR5 receptors expressed in HEK293 cells using whole-cell voltage-clamp recording (holding potential, $V_h = -70$ mV). **b**, ATPA (circles) and kainate (squares) concentration-response curves recorded in acutely isolated DRG neurons under whole-cell voltage-clamp conditions. **c**, ATPA (circles), kainate (squares) and AMPA (triangles) concentration-response curves recorded in acutely isolated Purkinje neurons under whole-cell voltage-clamp conditions. Data points for each graph represent the mean $\pm$ s.e.m. for at least 3 determinations from separate cells. Scale bars: **a**, 50 pA; **b**, **c**, 100 pA; and 10 s.

seen at high concentrations of LY294486 (for example, $33 \pm 2\%$ inhibition at $100 \mu\text{M}$, $n = 5$; Fig. 3).

The identification of the selective agonist and antagonist for GluR5 enables us to determine the functions of this receptor. We used ATPA and LY294486 to investigate the function of GluR5 in the hippocampus. We studied monosynaptic GABAergic inhibition²⁴,

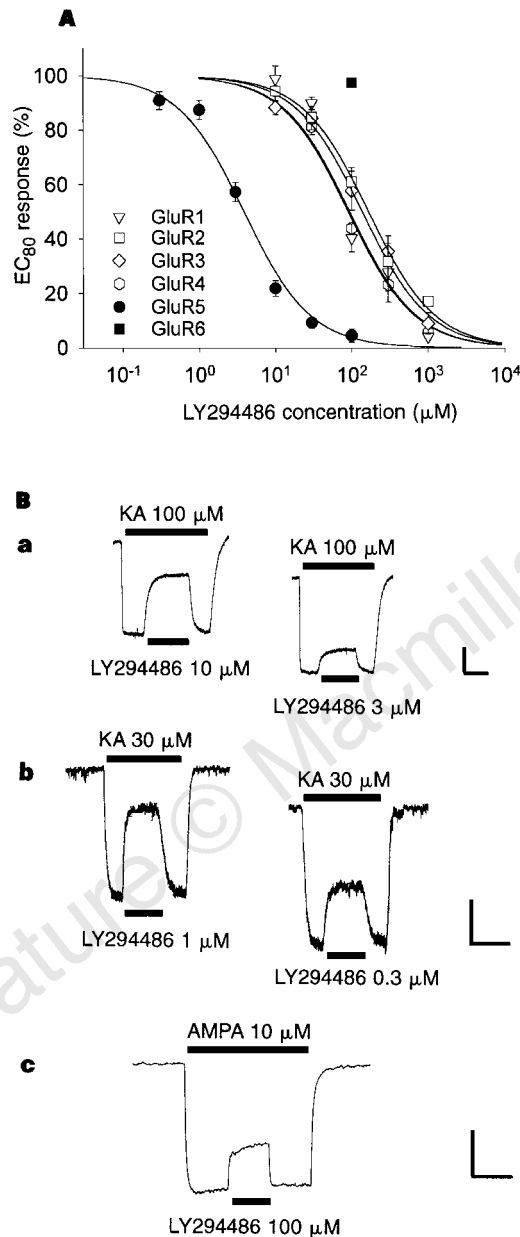


Figure 3 Antagonism of cloned human AMPA and kainate receptors by LY294486.

A, Concentration-dependent inhibition of agonist-induced inward currents evoked by glutamate ($100 \mu\text{M}$, GluR1 (open triangles), GluR2 (open squares), GluR3 (open diamonds) and GluR4 (open hexagons)) and kainate ($100 \mu\text{M}$, GluR5 (filled circles), $1 \mu\text{M}$, GluR6 (filled squares)) receptors. EC₅₀ values for glutamate at GluR1–4 were 36 ± 5 , 36 ± 5 ($n = 10$), 41 ± 4 ($n = 8$) and $94 \pm 3 \mu\text{M}$ ($n = 4$), respectively. EC₅₀ values for kainate at GluR5 and GluR6 were 16 ± 1 ($n = 6$) and $0.8 \pm 0.1 \mu\text{M}$ ($n = 7$), respectively. **B**, Sample current traces demonstrating the inhibition of agonist-induced currents by LY294486 at **a**, cloned human GluR5 expressed in HEK293 cells, scale bar denotes 2 s and 20 pA; **b**, rat DRG neurons, scale bar denotes 10 s and 50 pA; and **c**, rat cerebellar Purkinje neurons, scale bar denotes 1 s and 200 pA. Agonists were used at approximate EC₅₀ values, $100 \mu\text{M}$ kainate (KA) at GluR5, $30 \mu\text{M}$ KA at DRG neurons and $10 \mu\text{M}$ AMPA for cerebellar Purkinje neurons.

as a reduction in synaptic inhibition has been suggested to contribute to the epileptogenic actions of kainate^{13,15,16}. GYKI53655 (LY300168) was used to eliminate AMPA receptor-mediated synaptic transmission while preserving functional kainate receptors^{25–29}. Kainate ($5 \mu\text{M}$) reversibly depressed monosynaptically activated inhibitory postsynaptic potentials (IPSPs) (to between 25 and 77% of control, $n = 5$) without affecting the membrane potential of the postsynaptic neuron. ATPA ($10 \mu\text{M}$) had a similar effect on the monosynaptic IPSP (depressions to between 34 and 88% of control, $n = 16$), but caused a small hyperpolarization of the postsynaptic neuron ($2.7 \pm 0.3 \text{ mV}$, $n = 16$). The effects of both kainate and ATPA were abolished, in a reversible manner, by LY294486 (30 – $100 \mu\text{M}$) (Fig. 4B).

The effects of kainate and ATPA on the monosynaptic IPSP are likely to be presynaptic because both the GABA_A and the GABA_B receptor-mediated components of the IPSP were depressed to a

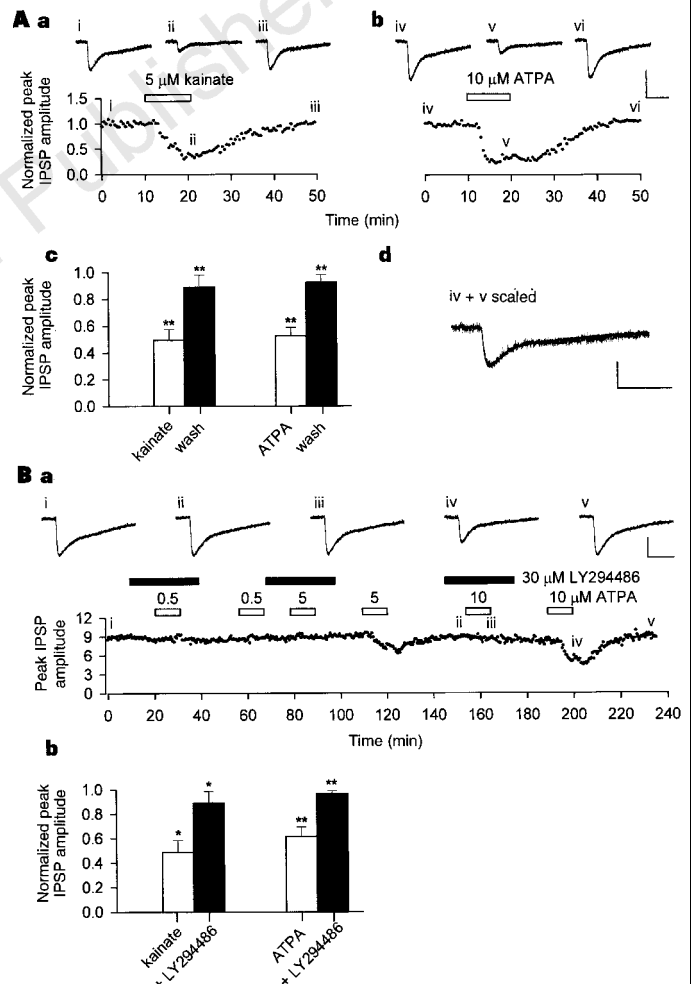


Figure 4 Activation of GluR5 depresses inhibitory synaptic transmission. **A**, A single example illustrates the depression of monosynaptic IPSPs by kainate (**a**) and ATPA (**b**). **c**, IPSP amplitude in the presence and following washout (for 30 min) for $5 \mu\text{M}$ kainate ($n = 5$) and $10 \mu\text{M}$ ATPA ($n = 16$). **d**, The maximally depressed response, in the presence of ATPA, is scaled and superimposed with a control response to show the lack of change in IPSP shape. **B**, **a**, A single example to show the effect of increasing concentrations of ATPA, in the presence and absence of LY294486. **b**, IPSP amplitude in the presence of either $5 \mu\text{M}$ kainate ($n = 4$) or $10 \mu\text{M}$ ATPA ($n = 5$). In each experiment the agonist was applied twice, once in the presence of LY294486 (30 – $100 \mu\text{M}$). Each graph plots the peak amplitude of individual monosynaptic IPSPs versus time, and the traces are averages of 4 successive responses obtained at the times indicated. Scale bars, 5 mV, 100 ms. * $P < 0.05$; ** $P < 0.01$.

similar extent (Fig. 4A). The alternative possibility that ATPA and kainate act postsynaptically to block both GABA_A and GABA_B receptor-mediated events seems unlikely, as this would require similar antagonism, and reversal of antagonism by LY294486, of two independent receptor/conductance mechanisms. A surprising observation was the small hyperpolarization induced by ATPA, which was never seen with kainate. The hyperpolarization persisted in the presence of either 50 μ M picrotoxin plus 10 μ M bicuculline (2.9 ± 0.6 mV, $n = 4$), or 10 μ M CGP55845A (2.8 ± 0.5 mV, $n = 4$), or all three antagonists (3.0 ± 0.8 mV, $n = 3$), and so is not due to direct or indirect activation of GABA receptors³⁰. Its antagonism by LY294486 suggests that it results from activation of GluR5-containing receptors.

The ability of ATPA to depress monosynaptic IPSPs was also unaffected by treatment with these GABA_A or GABA_B receptor antagonists. In one set of paired experiments, 10 μ M ATPA depressed monosynaptic IPSPs to $59 \pm 4\%$ and monosynaptic GABA_A receptor-mediated IPSPs to $59 \pm 10\%$ of control ($n = 4$). In a second set, it depressed monosynaptic IPSPs to $49 \pm 4\%$ and monosynaptic GABA_B receptor-mediated IPSPs to $48 \pm 3\%$ of control ($n = 4$). The finding that isolated GABA_A receptor-mediated IPSPs, recorded in the presence of 10 μ M CGP55845A, were depressed by ATPA excludes the possibility that ATPA depresses inhibition indirectly by depolarizing presynaptic terminals to release GABA, which then acts through presynaptic GABA_B receptors³⁰. Further experiments are required to determine the precise mechanism by which activation of GluR5-containing receptors depresses GABA-mediated synaptic inhibition.

Kainate receptor subunits (GluR5, GluR6, GluR7, KA-1 and KA-2) show a widespread and heterogeneous distribution throughout the central nervous system⁴⁻⁷. It has previously not been possible to address the function of any of these subunits. Several of the effects of kainate can be attributed to native kainate, rather than AMPA receptors, based on the use of doses of kainate subthreshold for AMPA-receptor activation or the use of AMPA-receptor-selective antagonists, such as GYKI53655. These effects include depolarization^{12,14}, depression of excitatory synaptic transmission²⁶, and, as we have shown, direct depression of inhibitory synaptic transmission. This latter effect can now be ascribed to receptors which are composed of, or contain, GluR5 subunits. The potent epileptogenic and neurotoxic actions of kainate are well known, but the mechanism of these actions has remained elusive. The finding that kainate receptor activation resulted in depression of monosynaptic inhibition in the absence of depolarization suggests that direct suppression of inhibition may be the primary factor in the epileptogenic actions of kainate, in which case GluR5 should provide a selective target for anti-epileptic drugs. In conclusion, the identification of a selective GluR5 kainate receptor agonist and a selective GluR5 kainate receptor antagonist should now enable the function of this receptor to be established. One possible role is the regulation of GABAergic synaptic transmission in the hippocampus. □

Methods

Ligand binding studies. Cell membranes were prepared from frozen HEK293 cells expressing either recombinant AMPA or kainate receptors by resuspending the cells in ice-cold distilled water, sonicating and centrifuging at 50,000g for 10 min. The membrane pellets were then washed in >100 vols of 50 mM Tris-HCl buffer, pH 7.5, and centrifuged to remove endogenous glutamate. Binding reactions were performed at 4°C for 60 min in a total volume of 250 μ l containing 50 μ l membrane suspension (100–150 μ g protein). For kainate receptor binding, the reaction mixture consisted of 150 μ l 50 mM Tris-HCl, pH 7.5, 25 μ l [³H]-KA (specific activity, 58.0 Ci mmol⁻¹; NEN, Dupont) and 25 μ l unlabelled competitor (10^{-11} – 10^{-3}). The final [³H]-KA concentrations used in the competitive inhibition experiments were as follows: 20 nM for GluR5 and GluR6; 10 nM for GluR6/KA-2; and 3 nM for GluR7 and KA-2. For AMPA receptor binding, 20 nM [³H]-AMPA (specific activity, 53.1 Ci mmol⁻¹; NEN,

Dupont) was used for each receptor subtype and 100 mM KSCN was added to the Tris-HCl buffer. After 60 min incubation, membranes were centrifuged at 50,000g for 20 min to separate bound from free ligand and the pellets were washed 3 times in cold assay buffer. Nonspecific binding was determined by incubation in the presence 10 μ M glutamate. All data were analysed using GRAFIT 2.0 software.

Electrophysiology. Whole-cell voltage-clamp recordings were made from isolated cells (HEK293 cells or primary cultured neurons obtained from rats 6–8 days old) using extracellular solutions composed of (in mM) 138, NaCl; 5, CaCl₂; 5, KCl; 1, MgCl₂; 10, HEPES; 10, glucose, pH to 7.4 with NaOH, osmolality 310 mOsm kg⁻¹ and intracellular solutions composed of (in mM): 140, CsCl; 1, MgCl₂; 14, diTris creatine phosphate; 50 U ml⁻¹, creatine phosphokinase; 10, HEPES and 15, BAPTA; pH to 7.15 with CsOH and osmolality of 295 mOsm kg⁻¹. For DRG neurons, intracellular solutions were composed of 125, CsMeSO₄; 15, CsCl; 5, CsBAPTA; 10, HEPES; 0.5, CaCl₂; 3, MgCl₂; 2, MgATP; pH 7.2 and osmolality 290 mOsm kg⁻¹. Experiments were performed at room temperature (20–22 °C) and recorded on either a List EPC-7 or Axopatch 1D amplifier. Drug application was through a series of perfusion lines to a multibarrelled applicator (Biologic Inc.) and exchange of solutions under the present conditions approximated 100 ms. Experiments using human GluR1–4 receptors expressed in HEK293 cells were performed in the presence of 30 μ M cyclothiazide. Cyclothiazide (100 μ M) was present for responses recorded in cerebellar Purkinje neuron experiments. Experiments using GluR5 and GluR6 receptors were performed following incubation of cells with 250 μ g ml⁻¹ concanavalin A for 10 min to remove desensitization²⁰. Curve fitting to the data points was based upon the equation $y = 100(D^n)/(D^n + EC_{50}^n)$ using a slope fixed to a value of 1, where D is the drug concentration. For antagonists, $EC_{50} = IC_{50}$. IC_{50} values were estimated from data obtained from a minimum of 4 separate cells. Intracellular recordings were obtained from the CA1 cell body region of hippocampal slices obtained from rats 5–8 weeks old as described²⁴. Monosynaptic IPSPs were obtained by stimulation in the stratum radiatum close (<0.5 mm) to the recording site, following blockade of AMPA receptors (using 50 μ M GYKI53655) and NMDA receptors (using 50 μ M D-2-amino-5-phosphonopentanoate plus, in some experiments, 5 μ M L-689,560). Data are presented as mean $\pm$ s.e.m.% control (relative to the 10-min baseline preceding drug administration) and were tested for significance using paired Student's *t*-tests. For each neuron, drug-induced effects were measured at the same membrane potential.

Received 9 April; accepted 17 July 1997.

- Watkins, J. C. & Evans, R. H. Excitatory amino acid transmitters. *Annu. Rev. Pharmacol. Toxicol.* **21**, 165–204 (1981).
- Collingridge, G. L. & Lester, R. A. J. Excitatory amino acid receptors in the vertebrate central nervous system. *Pharmacol. Rev.* **40**, 145–195 (1989).
- Monaghan, D. T., Bridges, R. J. & Cotman, C. W. The excitatory amino acid receptors: Their classes, pharmacology and distinct properties in the function of the central nervous system. *Annu. Rev. Pharmacol. Toxicol.* **29**, 365–402 (1989).
- Nakanishi, N. Molecular diversity of glutamate receptors and implications for brain function. *Science* **258**, 597–603 (1992).
- Seeburg, P. H. The molecular biology of mammalian glutamate receptor channels. *Trends Neurosci.* **14**, 297–303 (1993).
- Hollmann, M. & Heinemann, S. Cloned glutamate receptors. *Annu. Rev. Neurosci.* **17**, 31–108 (1994).
- Bettler, B. & Mülle, C. AMPA and kainate receptors. *Neuropharmacology* **34**, 123–139 (1995).
- Fletcher, E. J. & Lodge, D. New developments in the molecular pharmacology of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate and kainate receptors. *Pharmacol. Ther.* **70**, 65–89 (1996).
- Feldmeyer, D. & Cull-Candy, S. Elusive glutamate receptors. *Curr. Biol.* **4**, 82–84 (1994).
- Lerma, J., Morales, M., Vincente, M. A. & Herreras, O. Glutamate receptors of the kainate type and synaptic transmission. *Trends Neurosci.* **20**, 9–12 (1997).
- Bettler, B. *et al.* Cloning of a novel glutamate receptor subunit, GluR5: Expression in the nervous system during development. *Neuron* **5**, 583–595 (1990).
- Robinson, J. H. & Deadwyler, S. A. Kainic acid produces depolarization of CA3 pyramidal cells in the in vitro hippocampal slice. *Brain Res.* **221**, 117–127 (1981).
- Sloviter, R. S. & Damiano, B. P. On the relationship between kainic acid-induced epileptiform activity and hippocampal neuronal damage. *Neuropharmacology* **20**, 1003–1011 (1981).
- Westbrook, G. I. & Lothman, E. W. Cellular and synaptic basis of kainic acid-induced hippocampal epileptiform activity. *Brain Res.* **273**, 97–109 (1983).
- Kehl, S. J., McLennan, H. & Collingridge, G. L. Effects of folic and kainic acids on synaptic responses of hippocampal neurones. *Neuroscience* **11**, 111–124 (1984).
- Fisher, R. S. & Alger, B. E. Electrophysiological mechanisms of kainic acid-induced epileptiform activity in the rat hippocampal slice. *J. Neurosci.* **4**, 1312–1323 (1984).
- Ben-Ari, Y. & Gho, M. Long-lasting modification of the synaptic properties of rat CA3 hippocampal neurones induced by kainic acid. *J. Physiol. (Lond.)* **404**, 365–384 (1988).
- Krogsgaard-Larsen, P., Madsen, U., Ebert, B. & Hansen, J. J. in *Excitatory Amino Acid Receptors: Design of agonists and antagonists* (eds Krogsgaard-Larsen, P. & Hansen, J. J.) 34–55 (Ellis Horwood, London, 1994).
- Lauridsen, J., Honoré, T. & Krogsgaard-Larsen, P. Ibotenic acid analogues. Synthesis, molecular flexibility and in vitro activity of agonists and antagonists at central glutamic acid receptors. *J. Med. Chem.* **28**, 668–672 (1985).

20. Korczak, B. *et al.* cDNA cloning and functional properties of human glutamate receptor EAA3 (GluR5) in homomeric and heteromeric configuration. *Recept. Chann.* **3**, 41–49 (1995).
21. Huettner, J. E. Glutamate receptor channels in rat DRG neurons: Activation by kainate and quisqualate and blockade of desensitization by Con A. *Neuron* **5**, 255–266 (1990).
22. Partin, K. M., Patneau, D. K., Winters, C. A., Mayer, M. L. & Buonanno, A. Selective modulation of desensitization at AMPA versus kainate receptors by cyclothiazide and concanavalin A. *Neuron* **11**, 1069–1082 (1993).
23. Bleakman, D. *et al.* Pharmacological discrimination of GluR5 and GluR6 kainate receptor subtypes by (3S, 4aR, 6R, 8aR)-6-[2-(1H-tetrazole-5-yl)ethyl]decahydroisoquinoline-3-carboxylic acid. *Mol. Pharmacol.* **49**, 581–585 (1996).
24. Davies, S. N. & Collingridge, G. L. Role of excitatory amino acid receptors in synaptic transmission in area CA1 of rat hippocampus. *Proc. R. Soc. Lond. B* **236**, 373–384 (1989).
25. Paternain, A. V., Morales, M. & Lerma, J. Selective antagonism of AMPA receptors unmasks kainate receptor-mediated responses in hippocampal neurons. *Neuron* **14**, 185–189 (1995).
26. Chittajallu, R. *et al.* Regulation of glutamate release by presynaptic kainate receptors in the hippocampus. *Nature* **379**, 78–81 (1996).
27. Wilding, T. J. & Huettner, J. E. Differential antagonism of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid-preferring and kainate-preferring receptors by 2,3-benzodiazepines. *Mol. Pharmacol.* **47**, 582–587 (1995).
28. Partin, K. M. & Mayer, M. L. Negative allosteric modulation of wild-type and mutant AMPA receptors by GYKI53655. *Mol. Pharmacol.* **49**, 142–148 (1996).
29. Bleakman, D. *et al.* Activity of 2,3-benzodiazepines at native rat and recombinant human glutamate receptors in vitro: Stereospecificity and selectivity profiles. *Neuropharmacology* **35**, 1689–1702 (1996).
30. Davies, C. H., Pozza, M. F. & Collingridge, G. L. CGP55845A: A potent antagonist of GABA_A receptors in the CA1 region of the rat hippocampus. *Neuropharmacology* **32**, 1071–1073 (1993).

Acknowledgements. We thank M. Deverill for technical assistance. This work was supported by the MRC (G.L.C.) and the Wellcome Trust (V.R.J.C.).

Correspondence and requests for materials should be addressed to D.B.

Amyloidogenic role of cytokine TGF- β 1 in transgenic mice and in Alzheimer's disease

Tony Wyss-Coray^{*†}, Eliezer Masliah^{‡†}, Margaret Mallory[‡], Lisa McConlogue[§], Kelly Johnson-Wood[§], Carol Lin^{*} & Lennart Mucke^{*#}

^{*} Gladstone Molecular Neurobiology Program and Department of Neurology, University of California, San Francisco, California 94141-9100, USA

[#] Neuroscience Program, University of California, San Francisco, California 94143-0444, USA

[‡] Departments of Neurosciences and Pathology, University of California, San Diego, La Jolla, California 92093-0624, USA

[§] Athena Neurosciences Inc., 800 Gateway Boulevard, South San Francisco, California 94080, USA

[†] These authors contributed equally to this work.

Deposition of amyloid- β peptide in the central nervous system is a hallmark of Alzheimer's disease and a possible cause of neurodegeneration^{1–3}. The factors that initiate or promote deposition of amyloid- β peptide are not known. The transforming growth factor TGF- β 1 plays a central role in the response of the brain to injury^{4,5}, and increased TGF- β 1 has been found in the central nervous system of patients with Alzheimer's disease^{6–8}. Here we report that TGF- β 1 induces amyloid- β deposition in cerebral blood vessels and meninges of aged transgenic mice overexpressing this cytokine from astrocytes. Co-expression of TGF- β 1 in transgenic mice overexpressing amyloid-precursor protein, which develop Alzheimer's like pathology^{9–11}, accelerated the deposition of amyloid- β peptide. More TGF- β 1 messenger RNA was present in post-mortem brain tissue of Alzheimer's patients than in controls, the levels correlating strongly with amyloid- β deposition in the damaged cerebral blood vessels of patients with cerebral amyloid angiopathy. These results indicate that overexpression of TGF- β 1 may initiate or promote amyloidogenesis in Alzheimer's disease and in experimental models and so may be a risk factor for developing Alzheimer's disease.

In several lines of transgenic mice, astroglial overexpression of TGF- β 1 induced a strong upregulation of extracellular matrix proteins such as laminin, fibronectin and heparan sulphate proteoglycan in the central nervous system (CNS), particularly in the vicinity of

perivascular astrocytes expressing TGF- β 1 (ref. 12). This and other studies^{6,29} have raised the possibility that TGF- β 1 may be involved in amyloidogenesis. Because extracellular matrix proteins have been implicated in the deposition of amyloid- β ¹³, which may play a key role in the pathogenesis of Alzheimer's disease^{2,3}, we investigated whether constitutive astroglial overexpression of TGF- β 1 would result in the development of Alzheimer-like alterations in the CNS.

The animals analysed were from the previously described¹² low-expressing glial fibrillary acidic protein (GFAP)-TGF- β 1 transgenic lines 64 ($n = 36$) and 115 ($n = 33$). In both lines, transgene-derived bioactive TGF- β 1 is produced by brain astrocytes from birth¹². Although astroglial TGF- β 1 expression in these mice is widespread, it is most prominent in perivascular locations (Fig. 1). Staining of brain sections from 12–18-month-old TGF- β 1 mice with thioflavin S, which is widely used to detect β -pleated amyloid proteins^{14,15}, revealed prominent amyloid deposits in cerebral blood vessels (Fig. 2b, c) and meninges (Fig. 2b). No such deposits were seen in non-transgenic control mice (Fig. 2a). Staining of blood vessels in TGF- β 1 mice was strikingly similar to the staining of blood vessels seen in patients with Alzheimer's disease and cerebral amyloid angiopathy (CAA) (Fig. 2d). Semiquantitative assessment of cerebral thioflavin S-positive deposits in TGF- β 1 mice and controls confirmed the strong association between TGF- β 1 expression and cerebrovascular amyloid deposition: the mean thioflavin S score was

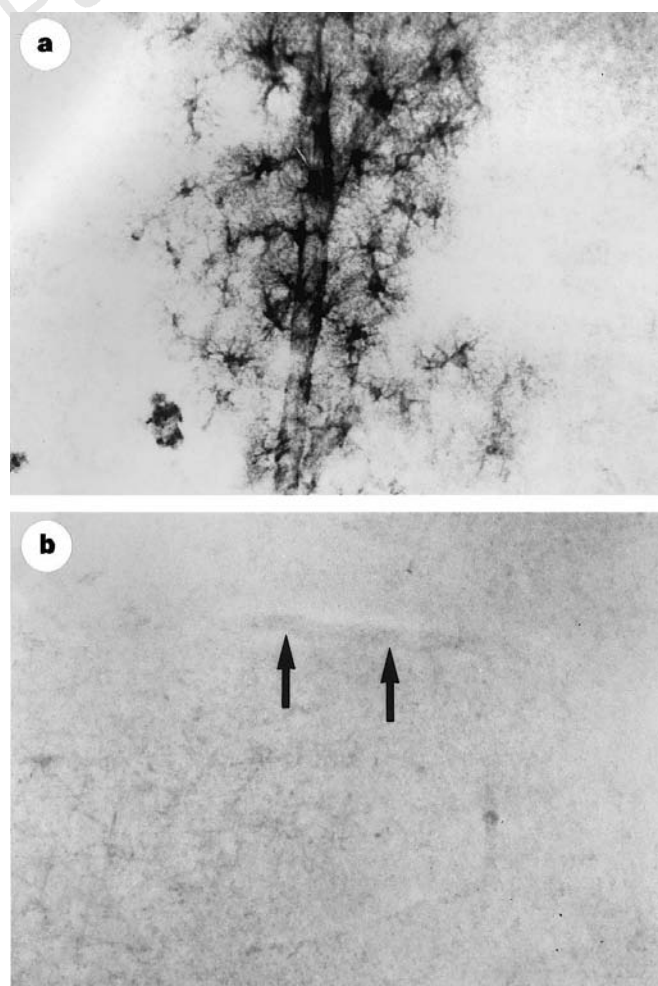


Figure 1 TGF- β 1 expression by perivascular astrocytes in TGF- β 1 mice. TGF- β 1 immunostaining of sagittal brain sections from a 16-month-old TGF- β 1 mouse (**a**) and a non-transgenic control (**b**) from line 64 with the transgene-specific TGF- β 1 antiserum G4. Note the labelling of transgenic astrocytes in the vicinity of the cerebral blood vessel. No transgene-derived TGF- β 1 was detectable in non-transgenic brains (**b**; arrows indicate the vessel). Original magnification: 135 $\times$.

2.1 ± 0.3 (range, 1–4) in TGF-β1 mice (n = 12) and 0.1 ± 0.1 (range, 0–1) in non-transgenic controls (n = 7). This amyloid deposition appeared to be both age- and dose-dependent, because amyloid deposits increased with age and, at a given age, were stronger and more abundant in homozygous transgenic mice showing more TGF-β1 expression than in heterozygous transgenic mice with less TGF-β1 expression (data not shown). Immunostain-

ing with different antibodies raised against human or rat amyloid-β labelled brain sections of aged TGF-β1 mice but not of non-transgenic controls in a pattern that was identical to that obtained with thioflavin S (data not shown). Additional studies are needed to determine the exact biochemical identity of the amyloid proteins deposited in these transgenic mice.

To test the ability of TGF-β1 to promote human amyloid-β deposition *in vivo*, we generated transgenic mice that express human amyloid-precursor protein (hAPP) and overproduce human amyloid-β (see Methods), and crossed them with TGF-β1 mice. High-expressor hAPP mice carrying amyloidogenic mutations produce high levels of human amyloid-β in the brain and develop prominent Alzheimer's-like pathology, including typical parenchymal amyloid plaques, in an age-related and brain-region-dependent way^{9–11,16,30}. Although some deposition of human amyloid-β also occurs around their cerebral blood vessels, cerebrovascular amyloidosis is not a prominent feature of these models, and it takes 6–8 months for heterozygous high-expressor hAPP mice to develop clear amyloid deposition¹¹. Therefore, if TGF-β1 promotes deposition of human amyloid-β, hAPP/TGF-β1 bigenic mice should develop amyloid deposits earlier. Furthermore, in view of the cerebrovascular amyloidosis observed in aged TGF-β1 mice, these deposits may first be detectable around cerebral blood vessels.

Table 1 Summary of amyloid-β/APP immunostaining

Antibody (reference) (final concentration)	Antigen	Immunostaining	
		hAPP mice (2–3-month-old)	hAPP/TGF-β1 mice (2–3-month-old)
10D5 mAb ¹⁷ (2 μg ml ⁻¹)	Aβ _{1–28} *	N	V, M, N
3D6 mAb ¹¹ (3 μg ml ⁻¹)	Aβ _{1–5}	No	V, M
2G3 mAb ¹¹ (10 μg ml ⁻¹)	Aβ _{33–40}	No	V, M
8E5 mAb ⁹ (3 μg ml ⁻¹)	hAPP _{444–592} †	N	N

mAb, monoclonal antibody; Aβ, amyloid-β peptide; hAPP, human amyloid precursor protein; V, vascular deposits; M, meningeal deposits; N, neuronal labelling. All antibodies were from Athena Neurosciences.

* Epitope: Aβ_{1–12}.

† Does not recognize amyloid-β.

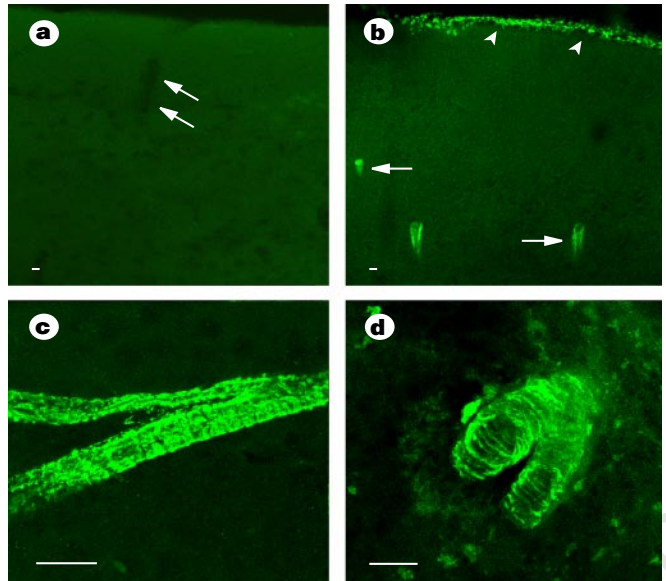


Figure 2 Thioflavin S-positive deposits in aged TGF-β1 mice and humans with Alzheimer's disease (AD). Thioflavin S staining of cortical brain sections from 16–18-month-old TGF-β1 mice (b, c), nontransgenic controls (a) (line 64), or an AD case with CAA (d; score of 3). Vascular staining of TGF-β1 mice (b, c) and AD cases with CAA (d) shows a similar appearance. Staining was also prominent in the meningeal layers of TGF-β1 mice (b). No staining was found in non-transgenic mice (a). Arrows, blood vessels; arrowheads, meninges. Scale bar, 20 μm.

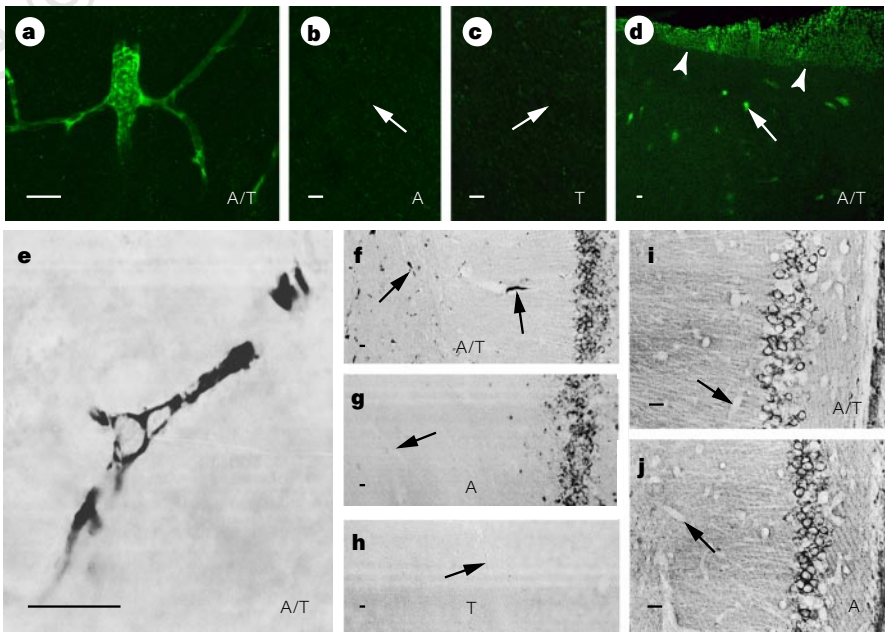


Figure 3 Accelerated development of cerebrovascular amyloid deposition in hAPP/TGF-β1 bigenic mice. Sagittal brain sections (a–e, neocortex; f–j, hippocampus) from 10-week-old hAPP/TGF-β1 (A/T) bigenic (a, d, e, f, i) and hAPP (A) (b, g, j) or TGF-β1 (T) (c, h) singly transgenic mice were stained with thioflavin S (d), monoclonal antibodies 3D6 (Aβ_{1–5}) (a–c), 2G3 (Aβ_{33–40}) (e), 10D5 (Aβ_{1–16}) (f–h), or the human APP holoprotein-specific antibody 8E5 (i, j). Brains

from bigenic mice showed prominent vascular (a, d–f) and meningeal (d) amyloid deposits but no vascular APP immunoreactivity (i). At this age, no vascular labelling was detected in corresponding brain sections from hAPP (b, g, j) or TGF-β1 (c, h) singly transgenic littermate controls. Arrows, vessels; arrowheads, meninges. Scale bar, 20 μm.

To test these ideas, we crossed heterozygous transgenic mice from PDGF-hAPP line H6 (see Methods) with heterozygous transgenic mice from GFAP-TGF- β 1 line 64. Co-expression of TGF- β 1 with hAPP/A β in bigenic mice resulted in cerebrovascular and meningeal (Fig. 3a, d–f) amyloid deposition at 2–3 months of age, whereas singly transgenic littermate controls representing either of the parental genotypes had no such deposits (Fig. 3b, c, g, h). The early amyloid deposits in bigenic mice stained with thioflavin S (Fig. 3d) as well as with amyloid- β antibodies 3D6, 2G3 and 10D5 (Fig. 3a, e, f). 3D6 detects the N terminus and 2G3 the C terminus of amyloid- β (Table 1). Antibody 10D5, which recognizes intact APP and amyloid- β ¹⁷, labelled both vascular structures and neurons in hAPP/TGF- β 1 bigenic mice (Fig. 3f), but labelled only neurons in hAPP singly transgenic mice (Fig. 3g). The human APP-specific antibody 8E5, which recognizes intact APP but not amyloid- β , labelled neurons with similar intensity in hAPP/TGF- β 1 bigenic and hAPP singly transgenic mice and did not label vascular structures (Fig. 3i, j). Measurements of hAPP and TGF- β 1 messenger RNA levels in brain by RNase protection assay revealed no significant differences between hAPP/TGF- β 1 bigenic mice and singly transgenic controls (data not shown). These results indicate that co-expression of TGF- β 1 in hAPP mice does not significantly alter hAPP expression. We conclude that TGF- β 1 promotes or induces the deposition of human amyloid- β *in vivo*.

The cerebrovascular thioflavin S and amyloid- β immunostaining in brains of TGF- β 1 mice closely resembles human CAA (Fig. 2), which is frequently found in brains of patients with Alzheimer's disease or Down's syndrome (for review, see refs 14, 15, 18). CAA is also associated with increased risk of stroke as well as cerebellar,

cerebral and subarachnoid hemorrhages^{14,18,19}. Our findings raise, to our knowledge for the first time, the possibility that TGF- β 1 may specifically cause the development of cerebrovascular amyloid deposition in Alzheimer's disease and related disorders. Until now, the aetiology of CAA has remained entirely obscure, and there has been no convenient animal model in which to study this prominent aspect of Alzheimer's pathology^{14,15}.

To test whether TGF- β 1 is involved in the development of CAA in humans, we examined post-mortem cortical tissues from Alzheimer's cases ($n = 12$) with different degrees of cerebrovascular amyloid deposition and from non-demented controls ($n = 5$). Strong TGF- β 1 immunoreactivity was seen around blood vessels in Alzheimer's cases with severe CAA, whereas no, or only faint, TGF- β 1 staining was observed in Alzheimer's cases without CAA (Fig. 4). As already shown^{6,8}, senile plaques and some neurons also have TGF- β 1 immunoreactivity (Fig. 4e).

To explore further the potential pathogenetic link between TGF- β 1 overproduction and the development of Alzheimer's disease, we compared TGF- β 1 mRNA levels in frontal cortex with CAA scores in cohorts of Alzheimer's patients and non-demented human controls (Fig. 5). Frontal cortex was chosen because this brain region typically develops prominent amyloid deposits in Alzheimer's disease. We found that Alzheimer's cases had threefold more TGF- β 1 mRNA than non-demented controls (Fig. 5a, b). In addition, Spearman's rank correlation analysis showed that TGF- β 1 mRNA levels correlated strongly with CAA scores (Fig. 5c). The results from our correlative analysis of human cases are consistent with our observations in TGF- β 1 transgenic mice.

Amino-acid substitutions in the different gene products APP, apolipoprotein E and presenilins strongly increase the risk of developing Alzheimer's disease^{2,3,20}; however, none of these changes appears to be required for development of the disease, underlining its multifactorial pathogenesis. Indirect evidence has been accumulating that inflammatory mediators may be important in the progression of this and other neurodegenerative diseases^{21,22}. Traumatic head injury is also a strong risk factor for Alzheimer's and can result in deposition of amyloid- β and the formation of neurofibrillary tangles at a young age^{23,24}. Studies of the PDGF-hAPP model⁹ indicate that, at least in mice, ageing is not associated with increased activity of the 'beta-secretase' pathway (which gives rise to

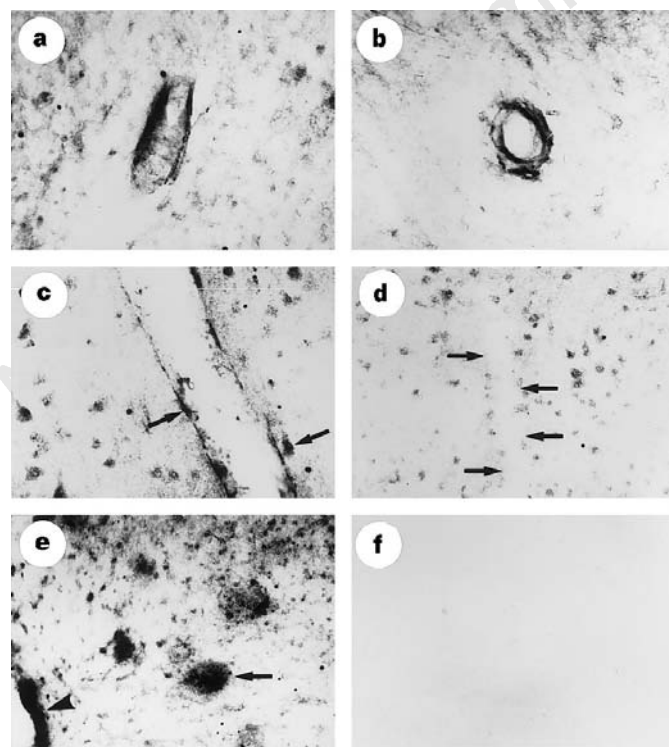


Figure 4 Vascular TGF- β 1 immunoreactivity in AD brains is associated with CAA. Immunostaining of frontal cortex sections from AD cases with CAA (a–c, e, f; score of 2) or without CAA (d) were stained with the TGF- β 1 antibody 100-NA. Strong TGF- β 1 immunoreactivity was detected along cerebral blood vessels (a–c) and meninges (e, arrowhead) in the AD patient with CAA, but not along cerebral blood vessels in the AD case without CAA (d, arrows). AD plaques were also labelled by the TGF- β 1 antibody (e, arrow). Preadsorption of the TGF- β 1 antibody with recombinant human TGF- β 1 blocked the immunostaining (f). Original magnification: 245 $\times$.

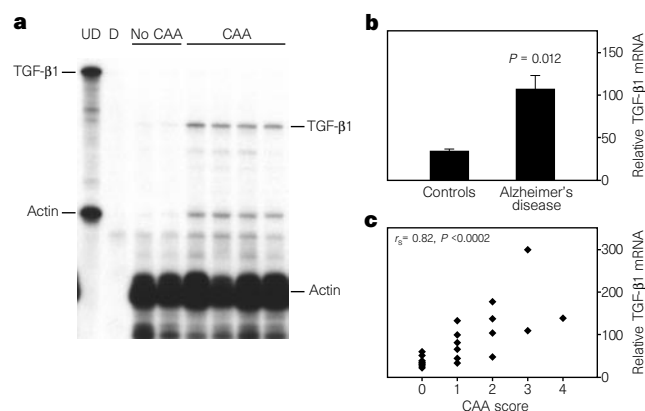


Figure 5 TGF- β 1 mRNA levels in frontal cortex of AD brains correlate with the degree of CAA. Relative TGF- β 1 mRNA levels were measured in frontal cortices of AD cases and controls by RPA with TGF- β 1 and actin antisense riboprobes. **a**, Results of RPA from a non-demented case, an AD case without CAA, and four AD cases with CAA (scores from left: 0, 0, 4, 3, 3, 2). UD, undigested probes; D, digested probes. **b**, Significantly higher TGF- β 1 mRNA levels were found in AD cases ($n = 15$; Blessed score ≥ 10) than in non-demented controls ($n = 7$; Blessed score = 0) by Mann-Whitney's unpaired two-tailed U test. **c**, Spearman's rank correlation analysis of TGF- β 1 mRNA levels and CAA scores from all 22 cases examined revealed a strong positive correlation.

the amyloid- β peptide) and that increased production of amyloid- β is not sufficient to induce amyloid deposition in certain brain regions¹¹. These findings strongly suggest that cofactors are critical for the development of Alzheimer's-related amyloidosis. We have shown that chronic glial overexpression of TGF- β 1 can result in the time- and dose-dependent development of thioflavin S-positive, amyloid- β -immunoreactive cerebrovascular and meningeal amyloid deposits *in vivo*, indicating that the injury-response factor TGF- β 1 may be important in Alzheimer's-related amyloidogenesis. TGF- β 1 may promote or mediate the deposition of amyloid in the CNS indirectly as a result of its capacity to induce amyloid- β -binding proteins such as extracellular matrix components^{4,12} and apolipoprotein E (T.W.-C. *et al.*, manuscript in preparation), which may function as seeds or chaperones for deposition of amyloid- β peptide^{13,25,26}. □

Methods

Mice and assessment of transgene products. Unless indicated otherwise, all transgenic mice were of Balb/c $\times$ SJL background and heterozygous for the respective transgene. Non-transgenic littermates served as controls. GFAP-TGF- β 1 lines 64 and 115 have been described¹². Heterozygous mice for line 64 express half the levels of cerebral TGF- β 1 mRNA expressed by heterozygous mice from line 115 (data not shown). High-level overexpression of TGF- β 1 results in the development of communicating hydrocephalus¹²; however, low-expressor mice such as those used here do not develop this complication. Human-APP transgenic mice were generated with the PDGF-hAPP transgene¹⁶, which induces robust Alzheimer's disease (AD)-like neuropathology when highly expressed in the CNS of transgenic mice^{9,10,11}. Cerebral hAPP/A β expression in the offspring from these lines was determined by RNase protection assay (RPA)¹⁶, western blot analysis⁹, and an ELISA to detect human amyloid- β ¹¹. Amyloid- β ELISA on whole-brain homogenates from 4–8-week-old transgenic mice from PDGF-hAPP line H6, the line used here, revealed slightly more cerebral amyloid- β production than in the previously described^{9,11} line 109 (data not shown).

Human AD cases. Frontal cortex (midfrontal gyrus) from cases with AD ($n = 15$; age 64–88 years) and from non-demented controls ($n = 7$; age 65–77 years) was obtained at autopsy as described¹⁶. Tissue blocks were either fixed in formalin or snap-frozen within 8 h after death.

Histology and immunocytochemistry. For thioflavin S staining, vibratome sections were made from human or murine brain tissue as described^{9,12,27} and air-dried overnight on superfrost slides (Fisher), fixed to the slides with 4% formaldehyde in 0.1 M phosphate buffer, and stained with a 1% thioflavin S solution for 8 min. Sections were rapidly washed once in 100% and twice in 80% ethanol/water, rinsed for 10 min with water, and mounted with Vectashield fluorescence mounting medium (Vector Labs). Sections were analysed by laser scanning confocal microscopy with a Bio-Rad MRC-1024 mounted on a Nikon Optiphot-2 microscope as described^{9,12,27}. The CAA score (for human brains) and the thioflavin S score (for murine brains) were determined semi-quantitatively by visual inspection of thioflavin S-stained brain sections: grade 0, no vessels affected; grade 1, occasional vessels affected; grades 2–4, multiple vessels affected mildly (grade 2), moderately (grade 3) or severely (grade 4) (for further details, see ref. 28). Two rabbit anti-TGF- β 1 antibodies were used for immunocytochemistry: G4 (ref. 12) and 100-NA (R&D Systems), each diluted 1:500. Antibodies directed against amyloid- β and APP are listed in Table 1.

Vibratome sections were stained by standard immunoperoxidase techniques using ABC kits from Vector Labs and 3,3'-diaminobenzidine plus hydrogen peroxide^{9,12}. Alternatively, sections were stained with monoclonal antibodies against amyloid- β (see above and Table 1), reacted with a Vectashield-labelled horse anti-mouse secondary antibody, and mounted with Vectashield fluorescence mounting medium. For preadsorption studies, recombinant human TGF- β (R&D Systems) was added to the TGF- β 1 antibody 100-NA solution to 10 ng ml⁻¹, and the antigen/antibody mixture was allowed to complex for 2 h at 37 °C. As a control, antibodies were incubated under identical conditions without antigen (no preadsorption). The mixture was then applied to the sections and processed as described for standard immunohistochemistry. Sections were analysed by standard light microscopy or laser-scanning confocal microscopy^{9,12,27}.

RNA extraction and analysis. High-quality total RNA was isolated from the frontal cortex of AD cases and controls as described¹⁶, and TGF- β 1 mRNA was determined by RPA as described¹². The TGF- β 1 riboprobe protects a 280- and a 125-nucleotide fragment of human TGF- β 1 mRNA (nucleotides 935–1,215 and 1,226–1,351 of human TGF- β 1 cDNA; GenBank accession no. X02812), but does not protect human TGF- β 2 or human TGF- β 3 transcripts. Levels of specific transcripts were compared by quantifying probe-specific signals with a phosphorimager (FUJI-BasIII), and β -actin signals were used to correct for differences in RNA content and loading¹⁶.

Statistical analysis. For histopathological analysis and for the RPA analysis of cortical TGF- β 1 mRNA, brain tissue samples were coded to blind investigators with respect to the genotype of the mice and to the diagnoses of the human cases examined. Analyses were done with Statview software (Abacus).

Received 14 March; accepted 22 July 1997.

- Selkoe, D. J. Alzheimer's disease: A central role for amyloid. *J. Neuropathol. Exp. Neurol.* **53**, 438–447 (1994).
- Lendon, C. L., Ashall, F. & Goate, A. M. Exploring the etiology of Alzheimer disease using molecular genetics. *J. Am. Med. Assoc.* **277**, 825–831 (1997).
- Hardy, J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci.* **20**, 154–159 (1997).
- Finch, C. E., Laping, N. J., Morgan, T. E., Nichols, N. R. & Pasinetti, G. M. TGF- β 1 is an organizer of responses to neurodegeneration. *J. Cell. Biochem.* **53**, 314–322 (1993).
- McCartney-Francis, N. L. & Wahl, S. M. Transforming growth factor β : A matter of life and death. *J. Leuk. Biol.* **55**, 401–409 (1994).
- van der Wal, E. A., Gómez-Pinilla, F. & Cotman, C. W. Transforming growth factor- β 1 is in plaques in Alzheimer and Down pathologies. *NeuroReport* **4**, 69–72 (1993).
- Chao, C. C. *et al.* Transforming growth factor β in Alzheimer's disease. *Clin. Diag. Lab. Immunol.* **1**, 109–110 (1994).
- Peress, N. S. & Perillo, E. Differential expression of TGF- β 1, 2 and 3 isoforms in Alzheimer's disease: A comparative immunohistochemical study with cerebral infarction, aged human and mouse control brains. *J. Neuropathol. Exp. Neurol.* **54**, 802–811 (1995).
- Games, D. *et al.* Alzheimer-type neuropathology in transgenic mice overexpressing V717F β -amyloid precursor protein. *Nature* **373**, 523–527 (1995).
- Masliah, E. *et al.* Comparison of neurodegenerative pathology in transgenic mice overexpressing V717F β -amyloid precursor protein and Alzheimer's disease. *J. Neurosci.* **16**, 5795–5811 (1996).
- Johnson-Wood, K. *et al.* Amyloid precursor protein processing and A β ₄₂ deposition in a transgenic mouse model of Alzheimer disease. *Proc. Natl Acad. Sci. USA* **94**, 1550–1555 (1997).
- Wyss-Coray, T. *et al.* Increased central nervous system production of extracellular matrix components and development of hydrocephalus in transgenic mice overexpressing transforming growth factor- β 1. *Am. J. Pathol.* **147**, 53–67 (1995).
- Fillit, H. & Leveugle, B. Disorders of the extracellular matrix and the pathogenesis of senile dementia of the Alzheimer's type. *Lab. Invest.* **72**, 249–253 (1995).
- Vinters, H. V. Cerebral amyloid angiopathy. A critical review. *Stroke* **18**, 311–324 (1987).
- Terry, R. D., Masliah, E., Hansen, L. A. in *Alzheimer's Disease* (eds Terry, R. D., Katzman, R. & Bick, K. L.) 179–196 (Raven, New York, 1994).
- Rockenstein, E. M. *et al.* Levels and alternative splicing of amyloid β protein precursor (APP) transcripts in brains of transgenic mice and humans with Alzheimer's disease. *J. Biol. Chem.* **270**, 28257–28267 (1995).
- Hyman, B. T., Tanzi, R. E., Marzloff, K., Barbour, R. & Schenk, D. Kunitz protease inhibitor-containing amyloid β protein precursor immunoreactivity in Alzheimer's disease. *J. Neuropathol. Exp. Neurol.* **51**, 76–83 (1992).
- Haan, J., Maat-Schieman, M. L. C. & Roos, R. A. C. Clinical aspects of cerebral amyloid angiopathy. *Dementia* **5**, 210–213 (1994).
- Itoh, Y., Yamada, M., Hayakawa, M., Otomo, E. & Miyatake, T. Cerebral amyloid angiopathy: A significant cause of cerebellar as well as lobar cerebral hemorrhage in the elderly. *J. Neurol. Sci.* **116**, 135–141 (1993).
- Strittmatter, W. J. & Roses, A. D. Apolipoprotein E and Alzheimer's disease. *Annu. Rev. Neurosci.* **19**, 53–77 (1996).
- McGeer, P. L. & McGeer, E. G. The inflammatory response system of brain: Implications for therapy of Alzheimer and other neurodegenerative diseases. *Brain Res. Rev.* **21**, 195–218 (1995).
- McGeer, P. L., Schuler, M. & McGeer, E. G. Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: A review of 17 epidemiologic studies. *Neurology* **47**, 425–432 (1996).
- Mayeux, R. *et al.* Genetic susceptibility and head injury as risk factors for Alzheimer's disease among community-dwelling elderly persons and their first-degree relatives. *Ann. Neurol.* **33**, 494–501 (1993).
- Nicoll, J. A. R., Roberts, G. W. & Graham, D. I. Apolipoprotein E ϵ 4 allele is associated with deposition of amyloid β -protein following head injury. *Nature Med.* **1**, 135–137 (1995).
- Wisniewski, T. & Frangione, B. Apolipoprotein E: a pathological chaperone protein in patients with cerebral and systemic amyloid. *Neurosci. Lett.* **135**, 235–238 (1992).
- Schmechel, D. E. *et al.* Increased amyloid β -peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 9649–9653 (1993).
- Masliah, E. *et al.* Spectrum of human immunodeficiency virus-associated neocortical damage. *Ann. Neurol.* **32**, 321–329 (1992).
- Olichney, J. M. *et al.* The apolipoprotein E ϵ 4 allele is associated with increased neuritic plaques and cerebral amyloid angiopathy in Alzheimer's disease and Lewy body variant. *Neurology* **47**, 190–196 (1996).
- Frautschi, S. A., Yang, F., Calderon, L. & Cole, G. M. Rodent models of Alzheimer's disease: rat A β infusion approaches to amyloid deposits. *Neurobiol. Aging* **17**, 311–321 (1996).
- Hsiao, K. *et al.* Correlative memory deficits, A β elevation, and amyloid plaques in transgenic mice. *Science* **274**, 99–102 (1996).

Acknowledgements. We thank F. Compagno and I. Samuels for technical assistance, D. Games for advice on the detection of amyloid deposits in transgenic mice, D. Schenk for amyloid- β antibodies, J. Rogers for post-mortem human brain tissue, and S. Ordway and G. Howard for editorial assistance. This work was supported by grants from the NIH and the Alzheimer's Association. T.W.-C. was partially supported by a fellowship from the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to T.W.-C. (e-mail: Tony_Wyss@quickmail.ucsf.edu).

An epithelial serine protease activates the amiloride-sensitive sodium channel

Véronique Vallet, Ahmed Chraïbi, Hans-Peter Gaeggeler, Jean-Daniel Horisberger & Bernard C. Rossier

Institut de Pharmacologie et de Toxicologie de l'Université, Rue du Bugnon 27, CH-1005 Lausanne, Switzerland

Sodium balance, and ultimately blood pressure and extracellular fluid volume, is maintained by precise regulation of the activity of the epithelial sodium channel (ENaC)¹⁻³. In a *Xenopus* kidney epithelial cell line (A6), exposure of the apical membrane to the protease inhibitor aprotinin reduces transepithelial sodium transport. Sodium-channel activity can be restored by subsequent exposure to the nonspecific protease trypsin. Using A6 cells and a functional complementation assay to detect increases in ENaC activity, we have cloned a 329-residue protein belonging to the serine protease family. We show that coexpression of this protein with ENaC in *Xenopus* oocytes increases the activity of the sodium channel by two- to threefold. This channel-activating protease (CAP1) is expressed in kidney, gut, lung, skin and ovary. Sequence analysis predicts that CAP1 is a secreted and/or glycosylphosphatidylinositol-anchored protein: ENaC activity would thus be regulated by the activity of a protease expressed at the surface of the same cell. This previously undiscovered mechanism for autocrine regulation may apply to other ion channels, in particular to members of the ENaC family that are present in neurons and epithelial cells.

The activity of ENaC is tightly controlled by hormones such as aldosterone and vasopressin, and also by intracellular and non-hormonal extracellular factors¹. As regards luminal factors, urinary proteases, such as kallikrein and urokinase, have been proposed to have a direct inhibitory effect on the sodium channel⁴; conversely, aprotinin, a serine-protease inhibitor, exerts an inhibitory effect on the short-circuit current in toad urinary bladder⁵, suggesting an activating role for serine proteases in the regulation of transepithelial sodium transport.

We have now analysed the effects of trypsin (a serine protease) and aprotinin on ENaC activity in A6 cells, a *Xenopus* kidney cell line expressing ENaC⁶. We found that overnight exposure of the apical membrane to aprotinin (28 µg ml⁻¹) resulted in a decrease in sodium transport (measured as the amiloride-sensitive short-circuit current, I_{sc} : aprotinin, $7.7 \pm 0.7 \mu A cm^{-2}$, $n = 29$; control, $15.2 \pm 1.3 \mu A cm^{-2}$, $n = 20$, $P < 0.001$). To test whether this effect of aprotinin was due to an inhibition of proteolytic activity or to some nonspecific action, we added trypsin in excess over aprotinin (5:1), which resulted in a significant increase (1.95 ± 0.06 -fold, $n = 14$) in I_{sc} (between the times indicated by the arrows in the middle panel of Fig. 1) compared with the control (1.41 ± 0.06 -fold, $n = 7$, $P < 0.001$; lower panel in Fig. 1). Results were similar with chymotrypsin (data not shown). Removal of aprotinin, whether or not trypsin was present, also induced a significant increase in I_{sc} (to 2.24 ± 0.20 -fold, $n = 12$) over the initial I_{sc} value in the presence of aprotinin, so that within a few minutes I_{sc} reached values comparable to those in control cells never exposed to aprotinin. These results indicate that the activity of the amiloride-sensitive sodium channel in A6 cells is modulated directly or indirectly by an endogenous protease that is active at the extracellular side of the apical membrane.

To isolate the protein(s) involved in the regulation of ENaC activity, we did a functional complementation assay in *Xenopus*

oocytes. An A6 cell complementary DNA library was screened by co-injection of cRNA from pools of cDNA together with the cRNA for the three $\alpha\beta\gamma$ subunits of *Xenopus* ENaC and measuring the amiloride-sensitive current³. We isolated a single cDNA clone whose coexpression with ENaC induced a ~ 3 -fold increase in the sodium current generated by ENaC. The 1,340-base-pair cDNA clone contains a 987-nucleotide open-reading frame encoding a 329-amino-acid protein (Fig. 2). Sequence comparison indicates that this protein shares homology with several members of the serine-protease family; we name this protein CAP1 (for channel-activating protease 1). The most closely related protein (53% identity) is a cloned human serine protease, prostaticin, found in semen⁷. Other serine proteases, such as trypsin, hepsin, kallikrein, tissue-plasminogen activator and urokinase, also show significant similarity to CAP1.

The hydropathy plot for the CAP1 protein shows that its amino- and carboxy-terminal regions each contain a stretch of hydrophobic residues (17 and 23 residues, respectively). A protein structure comprising a signal peptide at the N terminus and a hydrophobic domain at the C terminus, preceded by a characteristic signal⁸, is typical of glycosylphosphatidylinositol (GPI)-anchored proteins. As GPI anchoring is an apical targeting signal in polarized renal epithelial cells⁹, this property of CAP1 may provide the mechanism by which it is targeted to the apical cell surface of the highly polarized A6 cells, where the ENaC channel is located.

We next assessed the tissue distribution of CAP1 messenger RNA in *X. laevis* tissues by northern blot analysis (Fig. 3). A major mRNA species of about 1.5 kilobases (kb) is highly expressed in kidney, intestine, stomach, skin and lung, all epithelial tissues that express ENaC mRNA⁶. A longer exposure time allowed us to detect a faint signal in the ovary and in the eye. No signals were observed in liver, spleen, muscle or heart. Other transcripts with a higher molecular mass were detected in skin, kidney and stomach.

Next, the functional properties of CAP1 protein were studied by a two-electrode voltage-clamp method in *Xenopus* oocytes. As shown in Fig. 4a, coexpression of CAP1 together with the $\alpha\beta\gamma$ ENaC subunits led to a 2.8-fold increase in the amiloride-sensitive sodium current. CAP1 enhanced the activity of the rat ENaC channel

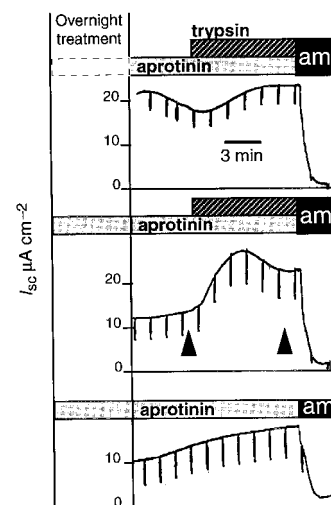


Figure 1 Effects of aprotinin and trypsin on sodium transport in A6 cells. In A6 cells not previously exposed to aprotinin (top trace), acute exposure to 20 µg ml⁻¹ aprotinin had a small inhibitory effect, and subsequent addition of trypsin (or removal of aprotinin; data not shown) induced only a small increase in I_{sc} . In contrast, in A6 cells exposed overnight to aprotinin and studied in the presence of aprotinin, addition of an excess of trypsin (200 µg ml⁻¹) induced a rapid increase in I_{sc} (middle trace); the same cells showed only a slow upward drift in the absence of trypsin (lower trace). am, Amiloride.

Cap1	1	MEPLPLSLFLTAUVHLEPSR
Prostasin		MAQKGVLPQGQI GAVAILTYLGLERSGT
Trypsin		MNPLILTFV
Hepsin		TSGFPCVDEGRPHTORLEPIVISYCDGPRGRF
Kallikrein		RCQFTYSLLPDCKEEKCKCFRLISMDGSPRIAYGTQSSSGYSLR
tPA		CSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMTLIGKYVTAQNPSAQALGLGKHNIC
Cap1	22	..SQEGV.....SRIVGGENATPGKFPWQVS
Prostasin		..GAECAEAPCG.....VAFQARIQSSSAVACQMPWQVS
Trypsin		AAALAAP.....FDDDDKIVGGVNCENSVFQVS
Hepsin		LAAICQDCGR.....KLVDRIVGGGRTSLRMPWQVS
Kallikrein		LCNTGDNVCT.....TKTSTRIVGGTNSWGEPWQVS
tPA		RNPDGDAKPWCHVLKNNRLTWEYCDVPSCSTCGLRQYSQPFQRIKGLFADIASHPWQAA
Cap1	47	RY...NGRHVCCASLISNXLTAACFPSPD.HLMSDYKVM LGVLOLEVTSESQ
Prostasin		I...TY...EGVHVCCGSLVSEQWVLSAAHCFPSE.HHKEAYEVKLCAGHOLDYSSEDAKV
Trypsin		L...N...SCVHFCCGSLINEQWVLSAGHCYKSR.....IQVRLCEHNIEVLEGNEQF
Hepsin		L...RY...DGAHLCCGSLISGDVLTAAHCFEERNVLSRWVFAO...AVAQASPHGLQ
Kallikrein		L...QVRLTAQRHLCCGSLTCHQWVLTAAHCF.DGLPLQDVWRISGLINTSDITKPTPF
tPA		IFAKHRRSPFGERFLCCGILISCTWLSAAHCFQERFF.PHHLTVLGRTRYRVVPGEEBQK
Cap1	100	LSLKEIITHFVS...DTSTGDUVALA...DPPATFSNVVQPIPLPDENVQFPI
Prostasin		STLRDIIFPHPSYLQ...EGSQGDIALQL...SRPITFSRYIRPICLPANASFPN
Trypsin		INAAKIIRHPPQDR.....KTLNNDIMLIK...SSRAVINARVSTISLETPAPP..AT
Hepsin		LGVAQAVVYHGGYLPFRDPNSSEENSNDIALVHL...SSPLPLTEYLOVCLPAAGQALVD
Kallikrein		SQIKETIITHQNY...KVSEGNHDIALLK...QAPLNYTEYQKICLPSKGDSTTI
tPA		FEVEKYLVHKEEDD...DTYDNDIALLOKSDSRCAQESVVRVCLLPADLQLED
Cap1	150	GMRQVTCWGNIQGVSPGSKTLQVGNVKILSRQTCNCLYHINPSSDSLGSVQDDMCA
Prostasin		GLHCTVTGWHVAPSVSLTPKPLQQLVPLISRETNCNLYNIDAKPEEPHFVQEDMCA
Trypsin		GRKCLZBQWGNASSCA.DYPDELQCLDAPVLSQAKCEASH.....PGKITSNMFCV
Hepsin		CKICTVTGWHNTQYYGQQ..AGVLQEARVPLISNDVNG.....ADFYGNQIKPKMCA
Kallikrein		YTNCTVTGWHGFSKERGEI..QNILOKVNIPLVTNEECQ.....KRYQDYKITQRMVCA
tPA		WTECEGEGYK.HEALSPFYSERKEAHVRLVPSRRCTSQHLLNRT.....YTDNMCA
Cap1	210	GSAAGS...VDACQGDGGGGLTCTV...NNQPYLAADVSWGDECGAFNRPGVYLLI
Prostasin		GVVEGG...KDACQGDGGGGLSCPV...EGLMYLTQIVSWGDACGARNRPGVYTLA
Trypsin		GVLEGG...KDSQGDGGGGLVYVNG...Q...LQGVSWGDGCAQKNRPGVYTKV
Hepsin		GVPEGG...IDACQGDGGGGLVVCEDSISRTPRRLCCIVSWGTGCGALAQKPGVYTKV
Kallikrein		GVKEGG...KDACQGDGGGGLVCK...HNGMRLVGLTISWGECCARREOPGVYTKV
tPA		GDTRSGGPQANLHDACQGDGGGGLVLC...LNDGRMTLVGLISWGLCGCQKQKVPGVYTKV
Cap1	260	SLSSWRSIDPSATVQYFTVDIPSDPQNNSGCVGADGQFYPNPNGA...SIFLVTFAL
Prostasin		SSVASWQS...VTELOPRVVPQTOESQPDNLGCSHLAFSSAPAGQLLRPILFLPL.GL
Trypsin		YNYVKWNTNTAANS.....
Hepsin		SDPREWTFQAITHSEASGMVTL.....
Kallikrein		AEVMDWILE...KTSQSDGKAQMSPA.....
tPA		TNGLDWIRDNMRP.....
Cap1	317	PFYWLTTYILSDF
Prostasin		ALGLLSPWLSEH*
Trypsin		
Hepsin		
Kallikrein		
tPA		

Figure 2 Comparison of the amino-acid sequence of *Xenopus* CAP1 with other serine proteases: human prostasin (1 to 363), human trypsin (1 to 247), human hepsin (114 to 417), human kallikrein (327 to 638) and human tissue-plasminogen activator (tPA) (208 to 562). Black and grey boxes indicate identical and similar residues, respectively. The catalytic triad is indicated by asterisks. The hydrophobic residues in N-terminal and C-terminal domains are underlined. Alignment was carried out using the Pileup program of the Genetic Computer Group sequence-analysis software package.

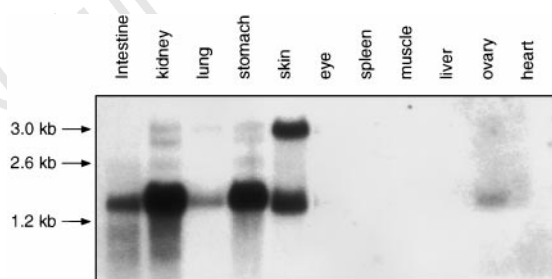


Figure 3 Northern blot analysis of CAP1 mRNA expression in *Xenopus laevis* tissues. Tissues shown on the left (intestine, kidney, lung, stomach, skin) were exposed for 7 h; those on the right half (eye, spleen, muscle, liver, ovary, heart) were exposed for 14 h.

(rENaC), as well as that of *Xenopus* channel (xENaC), even though the basal activity of rENaC was nearly 5-fold higher than that of xENaC ($2,023 \pm 261$ nA, $n = 32$, and 437 ± 87 nA, $n = 32$, respectively). To evaluate the specificity of CAP1, we tested the effect of CAP1 expression on ($\text{Na}^+ + \text{K}^+$) ATPase activity and found that it had no effect (activity measured at ($\text{Na}^+ + \text{K}^+$) pump current) in sodium-loaded oocytes¹⁰.

CAP1 could increase the sodium current mediated by ENaC by modifying the channel pore, thereby changing the single-channel

conductance and the ionic selectivity; changing the cell-surface expression of the channel; or changing the gating kinetics and thus increasing the open probability of the channel. First, the ion selectivity was assessed by replacing NaCl by LiCl or KCl in the oocyte's bathing solution. An amiloride-sensitive inward current was undetectable in the presence of K^+ and it increased by a factor of two when Na^+ was replaced by Li^+ (data not shown). The results were the same whether or not CAP1 was present. The single-channel conductance, measured by the patch-clamp method as an inward Li^+ current at positive pipette potentials, amounted to 7.9 ± 0.7 pS in oocytes expressing $\alpha\beta\gamma$ xENaC and CAP1 ($n = 4$). This value was essentially the same as the conductance of 7.3 ± 0.1 pS observed in oocytes expressing $\alpha\beta\gamma$ xENaC alone ($n = 12$). Next, the cell-surface expression of ENaC channels in oocytes was assessed by a binding assay using $\alpha\beta\gamma$ rENaC subunits tagged by a Flag epitope¹¹. CAP1 increased the amiloride-sensitive sodium current, but induced no significant change in the binding of anti-Flag antibodies (Fig. 4b). As the increase in I_{Na} induced by CAP1 cannot be explained by a change in the single-channel conductance nor in the cell-surface expression of ENaC, it must be due to an increase in the overall open probability of the channels present at the oocyte surface¹¹.

We have recently shown that the amiloride-sensitive sodium current in ENaC-injected oocytes can be enhanced ~2-fold by exposing the oocytes to $2 \mu\text{g ml}^{-1}$ trypsin¹² (see also Fig. 4c; compare lanes 1 and 2). CAP1 coexpression led to an increased basal sodium

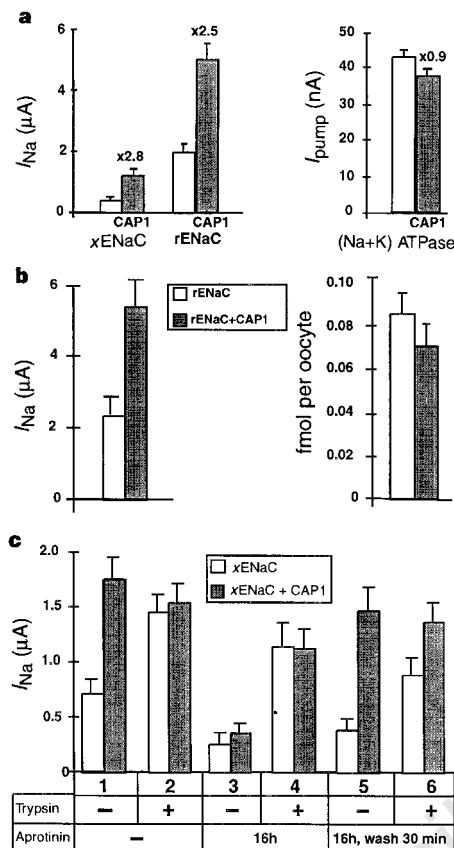


Figure 4 Functional effects of CAP1 expression on ENaC activity in *Xenopus* oocytes. Each bar represents the mean of four experiments on different batches of oocytes in which at least 8 measurements were done in each experimental group. **a**, Effect of CAP1 expression on ENaC and $Na^+ - K^+$ -ATPase activities. Macroscopic currents were recorded in *Xenopus* oocytes injected with $\alpha\beta\gamma$ *Xenopus* ENaC (xENaC) or $\alpha\beta\gamma$ rat ENaC (rENaC) subunit cRNAs in the presence of absence of CAP1 cRNA. Endogenous *Xenopus* $Na^+ - K^+$ -ATPase activity was measured in non-injected and CAP1-injected oocytes. The increase observed when CAP1 was coexpressed is indicated by the number above the relevant column. **b**, Effect of CAP1 expression on ENaC cell-surface expression. Flag-tagged ENaC expression at the surface of the oocytes was determined using an ^{125}I -labelled anti-Flag antibody and expressed as fmol per oocyte¹¹. The amiloride-sensitive sodium current (I_{Na}) and antibody binding were measured in the same oocytes (mean of 3 experiments, 10 measurements). **c**, Effects of trypsin and aprotinin on ENaC activity. Injected oocytes were incubated for 16 h in MBS buffer¹¹ in the presence or absence of 50 $\mu g\ ml^{-1}$ aprotinin. Groups (5 and 6) and aprotinin-incubated oocytes were washed for 30 min before measurements were made. The amiloride-sensitive sodium current was recorded, then oocytes were perfused for 2 min with 2 $\mu g\ ml^{-1}$ trypsin and the amiloride-sensitive sodium current measured again. All I_{Na} measurements were performed in the absence of aprotinin. Error bars represent the s.e.m.

current which could not be enhanced by subsequent exposure to trypsin. These results indicate that CAP1 and trypsin act through a common mechanism to enhance ENaC activity.

To test whether CAP1 acts by a proteolytic activity, oocytes were incubated for 16 h in the presence of aprotinin. Under these conditions, the basal amiloride-sensitive sodium current in CAP1-co-injected oocytes was no larger than that seen in ENaC-injected oocytes. However, this current was increased 3-fold by exposure to trypsin whether or not CAP1 was present (Fig. 4c, lanes 3 and 4). Thus, both of the effects of CAP1 (namely, activation of ENaC and the apparent abolition of the sensitivity of ENaC to trypsin) were prevented by incubation with aprotinin. These results

indicate that CAP1 protein may increase ENaC activity through the action of a serine protease.

The primary sequence of CAP1 protein predicts that it has an extracellular localization (that is, it is a secreted or membrane-anchored protein). When CAP1-injected oocytes were exposed overnight to aprotinin and the aprotinin then removed by washing for 30 min, the amiloride-sensitive sodium current recovered to a value close to the control and became insensitive to trypsin (Fig. 4c, lanes 5 and 6). Thus, as in A6 cells, the reversal of the aprotinin effect was fast, supporting the hypothesis that the action of CAP1 is exerted at sites located predominantly or entirely at the extracellular side of the plasma membrane. We tested whether CAP1 increases ENaC activity indirectly or by directly cleaving the extracellular loop of ENaC proteins, by specifically immunoprecipitating the α , β and γ Flag-tagged subunits¹¹: we found no evidence for cleavage (data not shown).

In addition to ENaC itself, which is mainly expressed in sodium-reabsorbing epithelia, channels encoded by this gene family include BNaC1 and BNaC2 (expressed in mammalian neurons¹³), and *mec-4* and *mec-10*, which encode components of a mechanosensory channel complex in *Caenorhabditis elegans* neurons¹⁴. An extracellular *C. elegans* protein, MEC-9, which shares Kunitz-like protease-inhibitor domains, has been cloned¹⁵; mutations in these domains altered the touch-sensitivity of the nematode. Moreover, MEC-9 seemed to be indirectly linked to the MEC-4/MEC-10 complex. These findings argue in favour of the idea that a general regulatory mechanism involving proteolytic events is implicated in the modulation of the activity of channels of the ENaC family.

ENaC plays a major part in the control of the blood pressure, as indicated by the discovery that mutations in ENaC genes are associated with hypertensive or hypotensive genetic diseases such as Liddle's syndrome¹⁶ or pseudohypoaldosteronism type 1 (ref. 17). CAP1 has been identified here as a protease involved in the regulation of ENaC activity. This finding suggests that CAP1 may be involved in the control of sodium balance and blood pressure, making it a candidate gene for the pathogenesis of hypertension. □

Methods

A6 cell culture and short-circuit current measurements. A subclone (2F3) of A6 cells from passages between 85–99 was cultured on Transwell permeable filters in modified HAM medium¹⁸ and treated with 300 nM aldosterone for 24 h before the experiment.

Construction of the A6 cells cDNA library. A6 cells were grown on filters and stimulated for 45 min with 300 nM aldosterone before mRNA extraction. The mRNA was size-fractionated by sucrose-gradient centrifugation. *Xenopus* oocytes were injected with $\alpha\beta\gamma$ ENaC cRNAs (1 ng) together with 50 ng of size-fractionated A6 mRNA, then the amiloride-sensitive sodium current was measured by a two-electrode voltage-clamp method. The 1–2-kb fraction was identified as inducing an increase in ENaC activity of ~40% over control (data not shown). A cDNA library was constructed³ from this 1–2-kb mRNA fraction, yielding ~30,000 independent clones.

Northern blot analysis. Isolation of mRNA and northern blot analysis were carried out as described⁶. Blots contained 10 μg per lane of poly(A)⁺ RNA from *Xenopus laevis* tissues. The membrane was hybridized with a full-length CAP1 cDNA probe.

Macroscopic current. Oocytes were injected with 1 ng cRNA for the $\alpha\beta\gamma$ xENaC subunits, together with 20–50 ng of cRNA from the A6 cDNA library or 1 ng of CAP1 cRNA. Electrophysiological measurements were made using a two-electrode voltage-clamp method as described¹¹. Results are the mean of at least 3 experiments, in each of which the amiloride-sensitive sodium current was measured in 6–10 individual oocytes. To test for the possible effect of CAP1 on $(Na^+ + K^+)$ pump activity, oocytes injected with CAP1 cRNA were Na^+ -loaded by overnight exposure in a K^+ -free solution and the Na^+ , K^+ -pump current was measured as the current activated by 10 mM K^+ at -50 mV (ref. 10).

Single-channel conductance measurements. Oocytes injected with $\alpha\beta\gamma$ xENaC alone or with $\alpha\beta\gamma$ xENaC and CAP1 were prepared for patch-clamp

experiments by manual removal of the vitelline membrane after exposure to a hypertonic solution¹². Cell-attached patches were obtained with pipettes (resistance of 10–20 MΩ) filled with a solution containing 80 mM Li⁺.

Received 30 May; accepted 8 August 1997.

- Garty, H. & Palmer, L. Epithelial sodium channels: function, structure, and regulation. *Physiol. Rev.* **77**, 359–396 (1997).
- Rossier, B. C. *Cum grano salis*: the epithelial sodium channel and the control of blood pressure. *J. Am. Soc. Nephrol.* **9**, 980–992 (1997).
- Canessa, C. M. et al. Amiloride-sensitive epithelial Na⁺ channel is made of three homologous subunits. *Nature* **367**, 463–467 (1994).
- Lewis, S. A. & Clausen, C. Urinary proteases degrade epithelial sodium channels. *J. Membr. Biol.* **122**, 77–88 (1991).
- Orce, G. G., Castillo, G. A. & Margolius, H. S. Inhibition of short-circuit current in toad urinary bladder by inhibitors of glandular kallikrein. *Am. J. Physiol.* **239**, F459–F465 (1980).
- Puotinen, A. et al. The highly selective low-conductance epithelial Na channel of *Xenopus laevis* A6 kidney cells. *Am. J. Physiol.* **38**, C188–C197 (1995).
- Yu, J. X., Chao, L. & Chao, J. Molecular cloning, tissue specific expression, and cellular localization of human prostasin mRNA. *J. Biol. Chem.* **270**, 13483–13489 (1995).
- Udenfriend, S. & Kodukula, K. How glycosyl-phosphatidylinositol-anchored membrane proteins are made. *Annu. Rev. Biochem.* **64**, 563–591 (1995).
- Lisanti, M. P., Caras, I. W., Davitz, M. A. & Rodriguez-Boulton, E. A. glycosphospholipid membrane anchor acts as an apical targeting signal in polarized epithelial cells. *J. Cell Biol.* **109**, 2145–2156 (1989).
- Horisberger, J. D., Jaunin, P., Good, P. J., Rossier, B. C. & Geering, K. Coexpression of β1 with putative β3 subunits results in functional Na–K-pumps in *Xenopus* oocytes. *Proc. Natl Acad. Sci. USA* **88**, 8397–8400 (1991).
- Firsov, D. et al. Cell surface expression of the epithelial Na channel and a mutant causing Liddle syndrome: a quantitative approach. *Proc. Natl Acad. Sci. USA* **93**, 15370–15375 (1996).
- Chraïbi, A., Firsov, D., Kharoubi Hess, S. & Horisberger, J.-D. Protease modulation of the activity of the epithelial sodium channel expressed in *Xenopus* oocytes. *J. Gen. Physiol.* (submitted).
- García-Anoveros, J., Derfler, B., Neville-Golden, J., Hyman, B. T. & Corey, D. P. BNaC1 and BNaC2 constitutes a new family of human neuronal sodium channels related to degenerins and epithelial sodium channels. *Proc. Natl Acad. Sci. USA* **94**, 1459–1464 (1997).
- Tavernarakis, N. & Driscoll, M. Molecular modeling of mechanotransduction in the nematode *Caenorhabditis elegans*. *Annu. Rev. Physiol.* **59**, 659–689 (1997).
- Du, H., Gu, G., William, C. M. & Chalfie, M. Extracellular proteins needed for *C. elegans* mechanosensation. *Neuron* **16**, 183–194 (1996).
- Shimkets, R. A. et al. Liddle's syndrome: Heritable human hypertension caused by mutations in the β subunit of the epithelial sodium channel. *Cell* **79**, 407–414 (1994).
- Chang, S. S. et al. Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypaldosteronism type 1. *Nature Genet.* **12**, 248–253 (1996).
- Handler, J. S. et al. Toad urinary bladder epithelial cells in culture: Maintenance of epithelial structure, sodium transport, and response to hormones. *Proc. Natl Acad. Sci. USA* **76**, 4151–4155 (1979).

Acknowledgements. We thank E. Hummler, K. Geering, M. Schapira and L. Schild for critically reading the manuscript. This work was supported by grants from the Swiss National Foundation for Scientific Research and Human Frontier Science Program; V.V. was supported by a fellowship from the European Molecular Biology Organization.

Correspondence and requests for materials should be addressed to B.C.R. (e-mail: Bernard.Rossier@ipham.unil.ch).

Protection from obesity-induced insulin resistance in mice lacking TNF-α function

K. Teoman Uysal^{*‡}, Sarah M. Wiesbrock^{*‡}, Michael W. Marino[†] & Gökhan S. Hotamisligil^{*}

^{*} Division of Biological Sciences and Department of Nutrition, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115, USA

[†] Ludwig Institute for Cancer Research, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

[‡] These authors contributed equally to this study.

Obesity is highly associated with insulin resistance and is the biggest risk factor for non-insulin-dependent diabetes mellitus^{1–3}. The molecular basis of this common syndrome, however, is poorly understood. It has been suggested that tumour necrosis factor (TNF)-α is a candidate mediator of insulin resistance in obesity, as it is overexpressed in the adipose tissues of rodents and humans^{4–10} and it blocks the action of insulin in cultured cells and whole animals^{10–14}. To investigate the role of TNF-α in obesity and insulin resistance, we have generated obese mice with a targeted null mutation in the gene encoding TNF-α and those encoding the two receptors for TNF-α. The absence of TNF-α

resulted in significantly improved insulin sensitivity in both diet-induced obesity and that resulting for the *ob/ob* model of obesity. The TNF-α-deficient obese mice had lower levels of circulating free fatty acids, and were protected from the obesity-related reduction in the insulin receptor signalling in muscle and fat tissues. These results indicate that TNF-α is an important mediator of insulin resistance in obesity through its effects on several important sites of insulin action.

It has been demonstrated previously that expression of TNF-α in adipose tissue is elevated in a variety of experimental obesity models^{4–6} and in obese humans^{7–10}, and might represent an important component of the link between obesity and insulin resistance^{4,15}. TNF-α blocks the action of insulin through its ability to inhibit insulin receptor tyrosine kinase activity^{10–14}, although other mechanisms, such as the quantitative regulation of glucose transporters, have also been proposed^{16,17}. To investigate the role of TNF-α in obesity-induced insulin resistance, we first generated obese mice that had no functional copy of the gene encoding TNF-α¹⁸. This was accomplished by placing mice homozygous for a targeted null mutation in the TNF-α gene (TNF-α^{-/-}) and their control littermates (TNF-α^{+/+}) on a high-fat (50% of the total calories in the form of fat) and high-caloric diet (5,286 kcal kg⁻¹, Bioserve, NJ). On this diet, both TNF-α^{-/-} and TNF-α^{+/+} mice developed marked obesity compared with mice kept on a standard rodent diet (Fig. 1a). In both the lean (standard diet) and obese (high-fat diet) groups, the total body weights of TNF-α^{-/-} and TNF-α^{+/+} mice were similar throughout the 16-week study. The total weight gain in the obese TNF-α^{+/+} animals on the high-fat diet tended to be slightly larger than that of the obese TNF-α^{-/-} mice (39.99 ± 0.9 versus 37.89 ± 1.12 g; Fig. 1a) but overall this difference was not significant. The average weights of the epididymal fat pad of the lean TNF-α^{+/+} and TNF-α^{-/-} mice were also similar (0.41 ± 0.09 versus 0.54 ± 0.06 g; Fig. 1b). In the obese group, fat pads of the TNF-α^{+/+} animals weighed ~22% more than those of the TNF-α^{-/-} mice (2.64 ± 0.15 versus 2.16 ± 0.14 g, *P* < 0.05). These results suggested that the absence of TNF-α did not have a significant effect on the development of dietary obesity, except for a potential small difference in adiposity. Further understanding of this issue requires detailed analysis of the body composition of these animals.

Because the effect of TNF-α in obesity is mainly on the action of insulin, we next investigated glucose homeostasis in TNF-α^{+/+} and TNF-α^{-/-} animals rendered obese by the high-fat diet. Measurement of fasting blood glucose levels demonstrated that all animals (lean and obese, TNF-α^{-/-} and TNF-α^{+/+}) remained euglycaemic throughout the study, although the lean TNF-α^{-/-} mice had the lowest fasting blood glucose levels (Fig. 2a). In contrast, fasting hyperinsulinaemia became apparent in obese TNF-α^{+/+} mice 4 weeks after the start of the high-fat diet and continued to increase in the following weeks (1.3 ± 0.3 ng ml⁻¹ at 8 weeks and 2.38 ± 0.6 ng ml⁻¹ at 12 weeks; Fig. 2b). However, the fasting

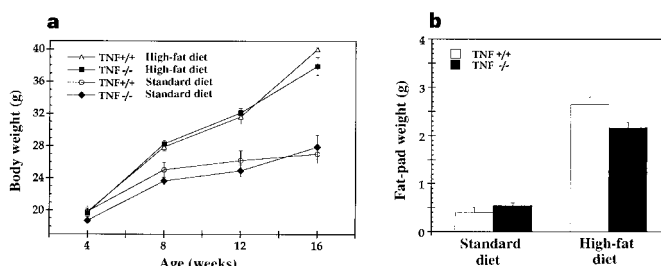


Figure 1 Growth curves and adiposity of TNF-α^{-/-} and TNF-α^{+/+} mice. **a**, Development of diet-induced obesity in TNF-α^{-/-} and TNF-α^{+/+} mice (*n* = 10 in each group). **b**, Epididymal fat-pad weights of TNF-α^{-/-} and TNF-α^{+/+} mice.

insulin concentrations in obese $\text{TNF-}\alpha^{-/-}$ mice were significantly lower ($0.4 \pm 0.07 \text{ ng ml}^{-1}$ at 8 weeks and 0.57 ng ml^{-1} at 12 weeks, $P < 0.001$) than those of the obese $\text{TNF-}\alpha^{+/+}$ animals (Fig. 2b), and were the same as those of the lean mice ($0.41 \pm 0.02 \text{ ng ml}^{-1}$ at 8 weeks and $0.627 \pm 0.2 \text{ ng ml}^{-1}$ at 12 weeks) throughout the study. At 12 weeks, the fasting insulin levels in the obese $\text{TNF-}\alpha^{+/+}$ group were increased ~fourfold compared with those of the obese $\text{TNF-}\alpha^{-/-}$ mice (Fig. 2b). The rise in serum insulin concentrations in the presence of euglycaemia indicates strongly that there is a compensatory response to the development of obesity-induced

insulin resistance in the wild-type obese mice. These results indicate that the mice deficient in $\text{TNF-}\alpha$ were protected from the development of obesity-induced insulin resistance that was observed in obese wild-type animals.

To investigate this possibility directly, we performed intraperitoneal insulin and glucose tolerance tests on these mice. The hypoglycaemic response to insulin was less in the obese $\text{TNF-}\alpha^{+/+}$ mice at 45–90 min than that in obese $\text{TNF-}\alpha^{-/-}$ animals (Fig. 2c). The intraperitoneal glucose tolerance tests also revealed a higher degree of hyperglycaemia in the obese $\text{TNF-}\alpha^{+/+}$ animals after 30–90 min than that in the obese $\text{TNF-}\alpha^{-/-}$ mice (Fig. 2d). Thus both of these tests indicated that there was marked insulin resistance in obese $\text{TNF-}\alpha^{+/+}$, but not in obese $\text{TNF-}\alpha^{-/-}$ mice. We did not observe any differences in the lean group between the $\text{TNF-}\alpha^{+/+}$ and $\text{TNF-}\alpha^{-/-}$ mice in insulin and glucose tolerance tests. These results demonstrate that the genetic absence of $\text{TNF-}\alpha$ can significantly

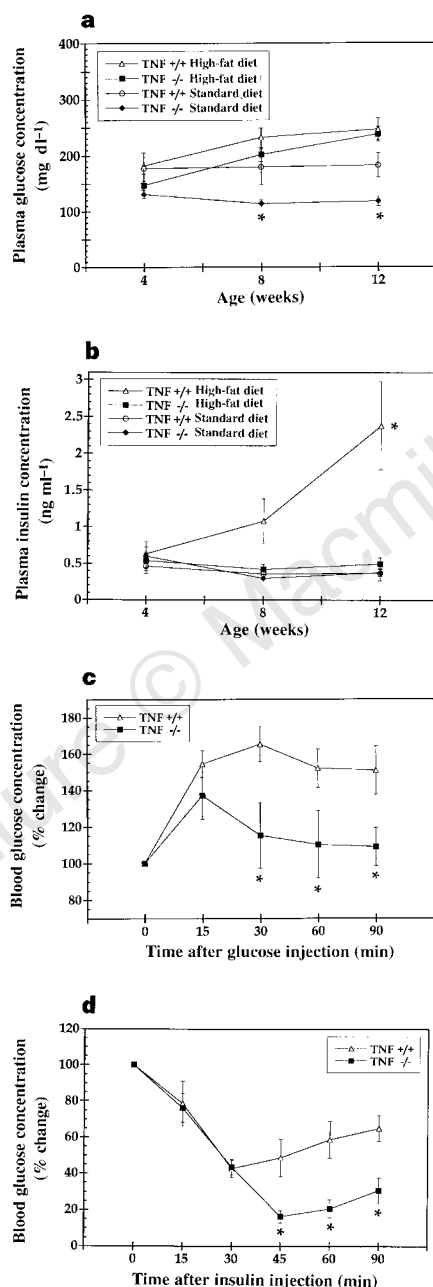


Figure 2 Measures of glucose homeostasis in $\text{TNF-}\alpha^{-/-}$ and $\text{TNF-}\alpha^{+/+}$ mice. **a, b**, Fasting glucose (**a**) and insulin (**b**) concentrations. **c, d**, Glucose (**c**) and insulin (**d**) tolerance tests. Asterisks indicate $P < 0.05$. Investigations of the dynamics of the responses to the tolerance tests were done by ANOVA repeated measures analysis (Statview 4.01, Abacus Concepts), and demonstrated statistically significant differences between the $\text{TNF-}\alpha^{-/-}$ and $\text{TNF-}\alpha^{+/+}$ mice in both tests ($P < 0.05$).

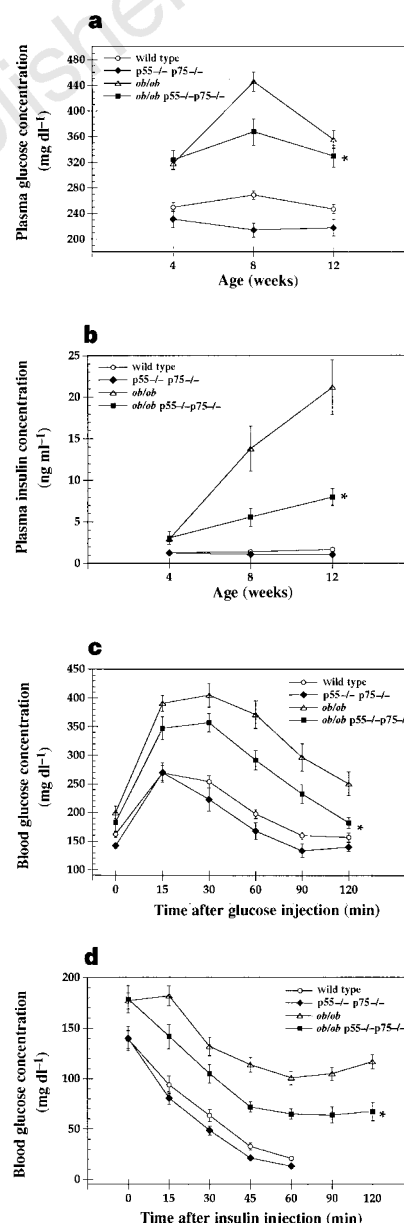


Figure 3 Measures of glucose homeostasis in ob/ob and $\text{ob/ob p55}^{-/-}\text{p75}^{-/-}$ mice. **a, b**, Fasting plasma glucose (**a**) and insulin (**b**) concentrations. **c, d**, Glucose (**c**) and insulin (**d**) tolerance tests. Mice studied: ob/ob ($n = 51$ in **a** and **b**, and 41 in **c** and **d**); $\text{ob/ob p55}^{-/-}\text{p75}^{-/-}$ ($n = 17$). Asterisks indicate $P < 0.05$.

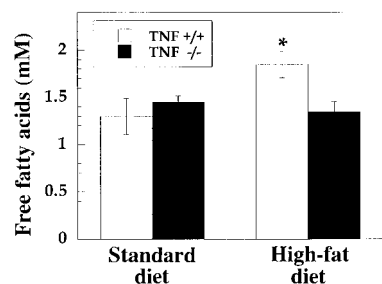


Figure 4 Circulating free fatty-acid levels in TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ mice. Asterisk indicates $P < 0.05$; two-tailed Student's t -test comparing free fatty-acid levels in TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ mice.

reduce development of insulin resistance associated with dietary obesity.

To test the importance of the TNF α -activated pathway of insulin resistance in the most severe, genetic model of obesity, we generated genetically obese (*ob/ob*) mice with targeted mutations in both p55 and p75 TNF receptors, effectively abolishing the signalling and function of TNF- α in these animals¹⁹. This experiment allowed us both to investigate the action of TNF- α in a different and more severe genetic model of obesity, and through an alternative strategy to block the action of TNF- α . As expected, the *ob/ob* mice developed early onset and severe obesity regardless of the TNF receptor allele they carried²⁰. There was no significant difference in body weights or body compositions between the obese animals lacking TNF- α function (*ob/ob p55^{-/-} p75^{-/-}*) and obese control animals (*ob/ob*) (unpublished data).

To determine the state of insulin sensitivity in these animals, we measured the fasting plasma glucose and insulin concentrations and performed insulin and glucose tolerance tests. As previously observed, *ob/ob* animals (wild-type at TNF receptor loci) developed a moderate and transient fasting hyperglycaemia by eight weeks of age (445.5 ± 15.3 mg dl⁻¹; Fig. 3a) that subsided at 12 weeks²⁰. In contrast, the increase in fasting blood glucose levels in the *ob/ob* mice lacking TNF- α function (*ob/ob p55^{-/-} p75^{-/-}*) was small and the hyperglycaemia was milder (367.2 ± 20.7 mg dl⁻¹ at 8 weeks of age). The *ob/ob* animals also displayed a severe and progressive hyperinsulinaemia during the course of the study (13.8 ± 2.7 and 21.2 ± 3.3 ng ml⁻¹ at 8 and 12 weeks of age, respectively; Fig. 3b). However, the *ob/ob* mice lacking TNF- α function (*ob/ob p55^{-/-} p75^{-/-}*) displayed significantly lower fasting plasma insulin levels throughout the study (5.6 ± 1.1 and 8.0 ± 1.0 ng ml⁻¹ at 8 and 12 weeks of age, respectively; Fig. 3b) than the *ob/ob* animals with functional TNF- α signalling.

We also determined insulin sensitivity in *ob/ob* mice that are wild type or mutant at the loci of both TNF receptors by performing intraperitoneal insulin and glucose tolerance tests (Fig. 3c, d). Both of these tests demonstrated significantly increased insulin sensitivity in the obese mice lacking TNF- α function (*ob/ob p55^{-/-} p75^{-/-}*) compared with obese controls (*ob/ob*). However, the *ob/ob* animals that are deficient in TNF receptors were still insulin resistant. This finding shows that interfering with TNF- α signalling through null mutations in both TNF receptors results in a significant but incomplete protection from the insulin resistance associated with the *ob/ob* phenotype. This indicates that several factors combine to result in insulin resistance in obesity, and that TNF- α is not solely responsible for non-insulin-dependent diabetes mellitus in mice.

After confirming the significant improvement in insulin sensitivity in mice lacking TNF- α function in both mild (diet-induced) and severe (genetically induced) forms of obesity, we examined the potential molecular mechanisms that might underlie the protection from insulin resistance in the absence of TNF- α function. For this analysis, we investigated the dietary model of obesity because it

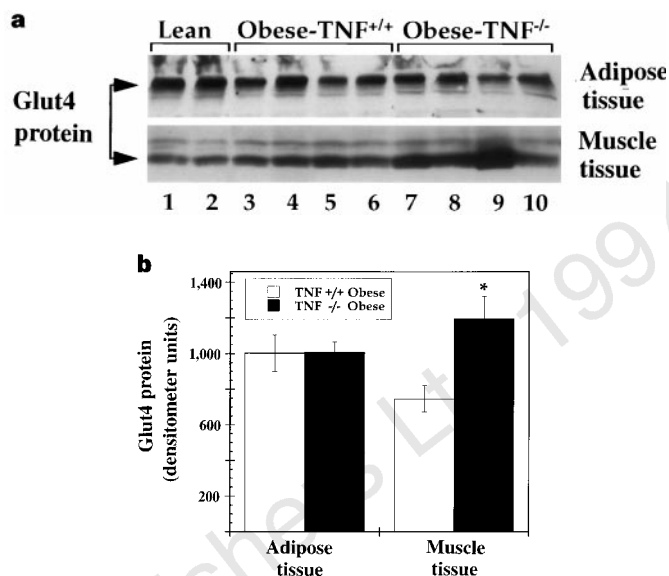


Figure 5 The levels of Glut4 protein in TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ mice. **a**, A representative immunoblot showing Glut4 protein level in fat and muscle tissues. **b**, Immunoblots are quantified by NIH Image 3.01 image-analysis software and presented as arbitrary units. Six mice were used in each analysis in two independent experiments. Asterisk indicates $P < 0.05$.

more closely resembles the level of obesity and insulin resistance seen in humans. We examined three potential sites of TNF- α action that might mediate insulin resistance: regulation of free fatty-acid levels; numbers of glucose transporters; and insulin receptor activity.

Elevated free fatty-acid levels in obesity are potential contributors to the development of insulin resistance¹. To investigate whether TNF- α deficiency influenced lipid metabolism, we measured serum free fatty-acid and triglyceride concentrations in TNF- $\alpha^{+/+}$ and TNF- $\alpha^{-/-}$ mice. Serum triglyceride levels in the lean TNF- $\alpha^{-/-}$ mice were lower (26%) than in the lean TNF- $\alpha^{+/+}$ animals (16.6 ± 1.4 versus 22.6 ± 3.3 mg dl⁻¹). However, no significant differences were observed in triglyceride levels between the obese TNF- $\alpha^{-/-}$ mice and the obese TNF- $\alpha^{+/+}$ animals (22.4 ± 3.8 versus 23.6 ± 3.4 mg dl⁻¹). As expected, the obese TNF- $\alpha^{+/+}$ animals had higher levels of circulating free fatty acids than the lean animals (1.84 ± 0.1 versus 1.3 ± 0.1 mM, $P < 0.05$; Fig. 4). In contrast, the free fatty-acid levels in the obese TNF- $\alpha^{-/-}$ mice were indistinguishable from those of the lean animals (1.34 ± 0.1 versus 1.3 ± 0.1 mM; Fig. 4). These results suggest that TNF- α contributes, either directly or indirectly, to the dyslipidaemia of obesity. This effect of TNF- α might be related to the increased insulin sensitivity in the obese TNF- $\alpha^{-/-}$ mice, as free fatty acids are believed to contribute to systemic insulin resistance in obesity.

Multiple sites of insulin signalling are known to be defective in obesity–diabetes syndromes²¹. These include the quantitative changes in insulin-sensitive glucose transporters and reduced signalling capacity of the insulin receptor in insulin-sensitive tissues²¹. It is important to understand the mechanistic basis of increased insulin sensitivity in the absence of TNF- α in obesity. In cultured adipocytes, hepatocytes, fibroblasts, muscle cells and myeloid cells, TNF- α inhibits the insulin-stimulated tyrosine kinase activity of the insulin receptor^{11–15}. In adipocytes and L6 myoblasts, TNF- α also downregulates the expression of the insulin-sensitive glucose transporter Glut4 (refs 4, 16, 22). Although pharmacological studies in rats have attributed most of the action of TNF- α to its effects on insulin-receptor signalling²³, the exact molecular basis of the TNF α -induced systemic insulin resistance *in vivo* has remained contro-

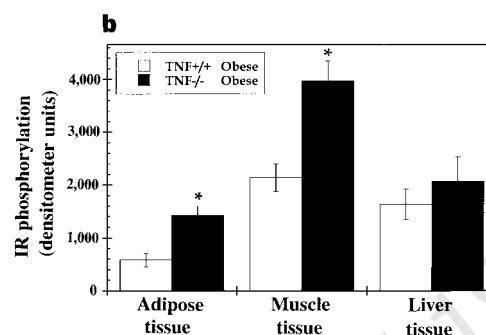
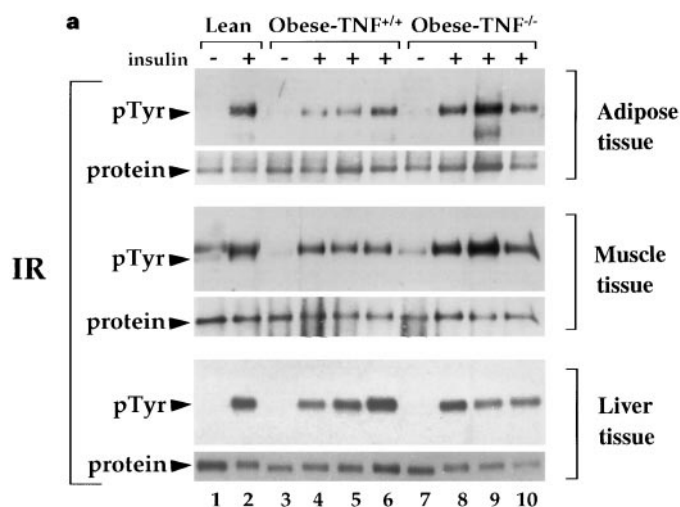


Figure 6 Levels of insulin-stimulated tyrosine phosphorylation of insulin receptor in fat, muscle and liver tissues of TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ mice. **a**, Representative immunoblot showing insulin-receptor (IR) tyrosine phosphorylation (pTyr) before (-) or after (+) insulin stimulation. **b**, Immunoblots are quantified by NIH Image 3.01 image-analysis software and presented as arbitrary units. Six mice were used in each analysis in three independent experiments. Asterisks indicate $P < 0.05$.

versal. To address these issues definitively, we examined the amount of Glut4 protein in the adipose and muscle tissues, and determined the insulin-stimulated tyrosine phosphorylation of the insulin receptor in fat, muscle and liver tissues of obese TNF- $\alpha^{+/+}$ and TNF- $\alpha^{-/-}$ mice.

In adipose tissue, levels of Glut4 protein in the obese mice were slightly lower than the lean controls (27%), but no significant difference was observed between the obese TNF- $\alpha^{+/+}$ and TNF- $\alpha^{-/-}$ animals (Fig. 5a, b). Because adipose tissue is the primary site of TNF- α expression in obese rodents⁴, this observation excludes the possibility that TNF- α is involved in the obesity-related quantitative regulation of Glut4 in adipose tissue. In muscle tissue, we did not observe a significant downregulation of Glut4 protein in TNF- $\alpha^{+/+}$ obese mice compared with the lean animals (Fig. 5a). Glut4 protein levels in muscle tissue were higher in TNF- $\alpha^{-/-}$ obese mice than in TNF- $\alpha^{+/+}$ animals (Fig. 5a, b). Although this result might suggest a role for TNF- α in the quantitative regulation of Glut4 in muscle, it is unlikely to explain the difference in insulin sensitivity seen between obese wild-type and TNF- α -deficient animals.

The binding of insulin to its receptor initiates a phosphorylation cascade that starts with the autophosphorylation of the insulin receptor on multiple tyrosine residues²⁴. This autophosphorylation is critical for the proper activation of the insulin-receptor tyrosine kinase and signalling events further downstream that mediate the biological actions of insulin²⁴. To determine the effects of TNF- α deficiency on the signalling capacity of the insulin receptor *in vivo*, we examined the ability of insulin to stimulate the tyrosine phosphorylation of the insulin-receptor β -chain in adipose, muscle and liver tissues of TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ animals. Following stimulation with insulin, phosphorylation of the 97k β subunit of the insulin receptor was visible in fat, muscle and liver tissues of all animals (Fig. 6). As expected, the insulin-stimulated autophosphorylation of the insulin receptor in obese TNF- $\alpha^{+/+}$ animals was lower than that of the lean animals in all tissues examined (70% in fat, 35% in muscle and 25% in liver; Fig. 6). There was a significant increase in autophosphorylation of the insulin receptor in the obese TNF- $\alpha^{-/-}$ animals compared with that of the TNF- $\alpha^{+/+}$ obese animals, which had levels approaching those observed in lean mice (Fig. 6). This increased insulin-stimulated phosphorylation of insulin receptor in the obese TNF- $\alpha^{-/-}$ animals was most significant in fat and muscle tissues (50% in fat and 46% in muscle, $P < 0.05$). The differences seen in insulin-receptor phosphorylation in liver were not statistically significant. In the lean group, the level of insulin-receptor phosphorylation was similar between the TNF- $\alpha^{-/-}$ and TNF- $\alpha^{+/+}$ mice. These results

demonstrate that, in the absence of TNF- α , the signalling capacity of the insulin receptor is significantly protected from obesity-induced downregulation in fat and muscle tissues.

We have provided clear evidence that the action of TNF- α is an important component of the link between obesity and insulin resistance in at least two different models (diet-induced and genetic, *ob/ob*) of murine obesity. As well as demonstrating the *in vivo* role of TNF- α in insulin resistance, our results also show that the action of TNF- α in obesity involves several potential targets that could influence systemic insulin action. First, the obese TNF- $\alpha^{-/-}$ mice have lower free fatty-acid levels than obese wild-type animals. This reduction in free fatty acids despite significant obesity might be the direct result of the loss of the lipolytic effects of TNF- α in adipose tissue, or alternatively might reflect the increased efficiency of insulin to suppress lipolysis in the absence of TNF- α ¹. Second, the obese TNF- $\alpha^{-/-}$ animals had higher levels of Glut4 protein in their muscle tissues. Finally, the obese TNF- $\alpha^{-/-}$ animals were spared from obesity-induced deficiencies in insulin-receptor signalling in fat and muscle tissues. This observation probably represents the most important effect of TNF- α in the generation of obesity-induced insulin resistance, owing to the critical importance of insulin-receptor signalling in generating the biological actions of insulin²⁴.

The extent of the role of TNF- α in the insulin resistance of dietary obesity appeared to be greater than its involvement in the *ob/ob* genetic model of obesity. This might simply be due to the extreme nature of the *ob/ob* phenotype compared to the dietary obesity. However, it is possible that the role of TNF- α is different as a result of the specific aetiology of obesity and the genetic background. Alternatively, TNF receptor deficiency might be distinct from TNF- α ligand deficiency, although existing data render this highly unlikely. Generation of additional cross-breeds with different obesity models, such as *db*, *tub* or *agouti*²⁵, should address these questions. □

Methods

Generation of *ob/ob* mice deficient in TNF receptors. Mice deficient in both TNF receptor 1 and TNF receptor 2 ($p55^{-/-} p75^{-/-}$; C57BL/6J129 mix) were back-crossed to C57BL/6 mice for three generations¹⁹, then intercrossed with *OB/ob* mice (Jackson Laboratories, Maine) to produce animals heterozygous at the $p55$, $p75$ and *ob* loci ($p55^{+/-}$, $p75^{+/-}$, *OB/ob*), which also involved two additional back-crosses to C57BL/6, the background strain of the *ob/ob* mutation. The resulting triple-heterozygote animals were cross-bred with each other to produce obese littermates with mutations in both TNF receptors (*ob/ob* $p55^{-/-} p75^{-/-}$) and with intact functional TNF receptors as controls (*ob/ob*).

Diet study and metabolic measurements. Male mice homozygous for a targeted null mutation at the TNF- α locus¹⁸ and their wild-type littermates (C57BL/6 $\times$ 129 genetic background) were housed in a barrier-free facility and given a high-fat, high-carbohydrate diet *ad libitum* (Diet F3282, Bioserve, NJ) at 4 weeks of age ($n = 10$), and were studied for the next 12 weeks. Identical groups of animals ($n = 5$) were given standard rodent chow to act as controls. Total body weights were measured weekly for 16 weeks, starting at 4 weeks of age. Blood samples were collected after a 6-h fast at 4, 8 and 12 weeks of age. Serum glucose concentrations were measured by using gluco-analyser blood glucose strips (Medisense). Serum insulin was measured with a monoclonal anti-rat insulin radioimmunoassay (Linco). The triglyceride and free fatty-acid levels in serum were determined (GPO Trinder, Sigma and Wako assays, respectively) using 12-week-old animals. Glucose and insulin tolerance tests were performed on conscious mice after a 6-h fast²⁶. Glucose tolerance tests were done by intraperitoneal administration of glucose (3 mg per g body weight) and measurement of blood glucose at 15, 30, 60 and 90 min in 18-week-old mice. Insulin tolerance tests were done similarly, except for the injection of human insulin (1 IU per kg; Eli Lilly) and an additional glucose measurement at 45 min. Fasting plasma glucose and insulin concentration and insulin sensitivity were also determined as described above for *ob/ob* and *ob/ob* p55^{-/-} p75^{-/-} mice.

In vivo insulin-stimulated insulin-receptor phosphorylation. After an overnight fast, mice were anaesthetized by intraperitoneal administration of 10 mg per kg xylazine and 100 mg per kg ketamine. The abdominal cavity was opened, and 25 mIU per kg insulin (Eli Lilly) or an equal volume of vehicle were administered through the portal vein. Liver, adipose (inguinal fat pads) and muscle (hindlimb) tissues were collected 120 s after the injection and immediately stored in liquid nitrogen. Protein extracts from the tissue samples were prepared as described²³.

Immunoprecipitation and immunoblotting. Total protein extract (750 μ g) was immunoprecipitated for 5 h at 4 °C by adding 1 μ g ml⁻¹ rabbit anti-insulin receptor antibody (gift from B. Cheatham and C. R. Kahn). Immune complexes were collected, washed and electrophoresed as described²³. After electrophoresis, the proteins were transferred to nitrocellulose membranes and protein immunoblot analysis was performed using a 1:2,000 dilution of a monoclonal anti-phosphotyrosine (gift from T. Roberts) or a 1:500 dilution of the polyclonal anti-insulin receptor as the primary antibody (gift from B. Cheatham and C. R. Kahn), followed by horseradish peroxidase-conjugated anti-mouse or anti-rabbit IgG antibodies (Promega) for detection. For Glut4 immunoblots, 100 μ g of protein extract was used and immunoblot analysis was performed by both a rabbit polyclonal (gift from B. Kahn) and a mouse monoclonal (gift from P. Pilch) anti-Glut 4 antibody, as described above.

Received 3 June; accepted 15 August 1997.

1. Boden, G. Role of fatty acids in the pathogenesis of insulin resistance and NIDDM. *Diabetes* **45**, 3–10 (1997).
2. Carey, V. J. et al. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women – the nurses health study. *Am. J. Epidemiol.* **145**, 614–619 (1997).
3. Edelstein, S. L. et al. Predictors of progression from impaired glucose tolerance to NIDDM. *Diabetes* **46**, 701–710 (1997).
4. Hotamisligil, G. S., Shargill, N. S. & Spiegelman, B. M. Adipose expression of tumor necrosis factor- α : direct role in obesity-linked insulin resistance. *Science* **259**, 87–91 (1993).
5. Hofmann, C. et al. Altered gene expression for tumor necrosis factor- α and its receptors during drug and dietary modulation of insulin resistance. *Endocrinology* **134**, 264–270 (1994).
6. Hamann, A. et al. Characterization of insulin resistance and NIDDM in transgenic mice with reduced brown fat. *Diabetes* **44**, 1266–1273 (1995).
7. Hotamisligil, G. S., Arner, P., Caro, J. F., Atkinson, R. L. & Spiegelman, B. M. Increased adipose expression of tumor necrosis factor- α in human obesity and insulin resistance. *J. Clin. Invest.* **95**, 2409–2415 (1995).
8. Kern, P. A. et al. The expression of tumor necrosis factor in adipose tissue: regulation by obesity, weight loss, and relationship to lipoprotein lipase. *J. Clin. Invest.* **95**, 2111–2119 (1995).
9. Saghizadeh, M., Ong, J. M., Garvey, W. T., Henry, R. R. & Kern, P. A. The expression of TNF α by human muscle: relationship to insulin resistance. *J. Clin. Invest.* **97**, 1111–1116 (1996).
10. Hotamisligil, G. S., Arner, P., Caro, J. F., Atkinson, R. L. & Spiegelman, B. M. Differential regulation of the p80 TNF receptor in human obesity and insulin resistance. *Diabetes* **46**, 451–455 (1997).
11. Feinstein, R., Kanety, H., Papa, M. Z., Lunenfeld, B. & Karasik, A. Tumor necrosis factor- α suppresses insulin-induced tyrosine phosphorylation of insulin receptor and its substrates. *J. Biol. Chem.* **268**, 26055–26058 (1993).
12. Hotamisligil, G. S., Murray, D. L., Choy, L. N. & Spiegelman, B. M. TNF- α inhibits signaling from insulin receptor. *Proc. Natl Acad. Sci. USA* **91**, 4854–4858 (1994).
13. Kanety, H., Feinstein, R., Papa, M. Z., Hemi, R. & Karasik, A. Tumor necrosis factor- α -induced phosphorylation of insulin receptor substrate-1 (IRS-1). *J. Biol. Chem.* **270**, 23780–23784 (1995).
14. Kroder, G. et al. Tumor necrosis factor- α and hyperglycemia-induced insulin resistance: evidence for different mechanisms and different effects on insulin signaling. *J. Clin. Invest.* **97**, 1471–1477 (1996).
15. Hotamisligil, G. S. et al. IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF- α - and obesity-induced insulin resistance. *Science* **271**, 665–668 (1996).

16. Stephens, J. M. & Pekala, P. H. Transcriptional repression of the GLUT4 and C/EBP genes in 3T3-L1 adipocytes by tumor necrosis factor- α . *J. Biol. Chem.* **266**, 21839–21845 (1991).
17. Stephens, J. M., Lee, J. L. & Pilch, P. F. Tumor necrosis factor- α -induced insulin resistance in 3T3-L1 adipocytes is accompanied by a loss of insulin receptor substrate-1 and glut4 expression without a loss of insulin receptor-mediated signal transduction. *J. Biol. Chem.* **272**, 971–976 (1997).
18. Marino, M. W. et al. Characterization of tumor necrosis factor deficient mice. *Proc. Natl Acad. Sci. USA* **94**, 8093–8098 (1997).
19. Bruce, A. J. et al. Altered neuronal and microglial responses to excitotoxic and ischemic brain injury in mice lacking TNF receptors. *Nature Med.* **2**, 788–794 (1996).
20. Coleman, D. L. in *The Mouse in Biomedical Research* (ed. Foster, M.) 125–132 (Academic, London, 1982).
21. Olefsky, J. M. & Molina, J. M. in *Diabetes Mellitus* (eds Rifkin, H. & Porte, D. J.) 121–153 (Elsevier, New York, 1990).
22. Cornelius, P., Lee, M. D., Marlowe, M. & Pekala, P. H. Monokine regulation of glucose transporter mRNA in L6 myotubes. *Biochem. Biophys. Res. Commun.* **165**, 429–436 (1989).
23. Hotamisligil, G. S., Budavari, A., Murray, D. L. & Spiegelman, B. M. Reduced tyrosine kinase activity of the insulin receptor in obesity-diabetes: central role of tumor necrosis factor- α . *J. Clin. Invest.* **94**, 1543–1549 (1994).
24. White, M. F. & Kahn, C. R. The insulin signaling system. *J. Biol. Chem.* **269**, 1–4 (1994).
25. Fiedorek, F. T. in *Diabetes Mellitus: a Fundamental and Clinical Text* (eds LeRoith, D., Taylor, S. I. & Olefsky, J. M.) 604–618 (Lippincott-Raven, Philadelphia, 1996).
26. Hotamisligil, G. S. et al. Uncoupling of obesity from insulin resistance through a targeted mutation in ap2, the adipocyte fatty acid binding protein. *Science* **274**, 1377–1379 (1996).

Acknowledgements. We thank J. Perschon for generation of the TNF receptor-deficient mice, and members of the Hotamisligil laboratory for discussions and support. This work is in part supported by grants from the NIH and American Diabetes Foundation (G.S.H.).

Correspondence and requests for materials should be addressed to G.S.H. (e-mail: ghotamis@hsph.harvard.edu).

Control of compartmental affinity boundaries by Hedgehog

Isabel Rodriguez* & Konrad Basler

Zoologisches Institut der Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

In *Drosophila*, each segmental primordium is subdivided into two cell populations, the anterior (A) and posterior (P) compartments by the selective activity of the transcription factor Engrailed (En) in P cells^{1–4}. Under En control, P cells secrete, but cannot respond to, the signalling protein Hedgehog (Hh)^{5–7}. In contrast, and by default, A cells are programmed to respond to Hh by expressing other signalling molecules, such as Decapentaplegic (Dpp) and Wingless (Wg), which organize growth and patterning in both compartments^{8,9}. Cells of the A and P compartments do not intermix, apparently as a consequence of their having distinct cell affinities that cause them to maximize contact with cells of the same compartment, while minimizing contact with cells from the other compartment¹⁰. This failure to mix has previously been ascribed to an autonomous and direct role for En in specifying a P, as opposed to an A, cell affinity^{3,11–13}. However, an alternative hypothesis is that Hh secreted by P cells induces A cells to acquire a distinct cell affinity, ensuring that a stable ‘affinity boundary’ forms wherever P and A cells meet. Here we show that the affinity boundary that segregates A and P cells into adjacent but immiscible cell populations is to a large extent a consequence of local Hh signalling, rather than a reflection of an intrinsic affinity difference between A and P cells.

To distinguish between these two hypotheses we used a mutation in the gene *smoothened* (*smo*)¹⁴, which encodes an essential component of the Hh signal-transduction pathway^{15,16}, to block the ability of A-compartment cells to receive and transduce Hh¹⁷. If distinct A and P cell affinities are specified autonomously by the state of En expression (‘off’ for A and ‘on’ for P), then anterior cells should retain their A-compartment affinity and sort out from P cells, even if their ability to respond to Hh is blocked by the loss of

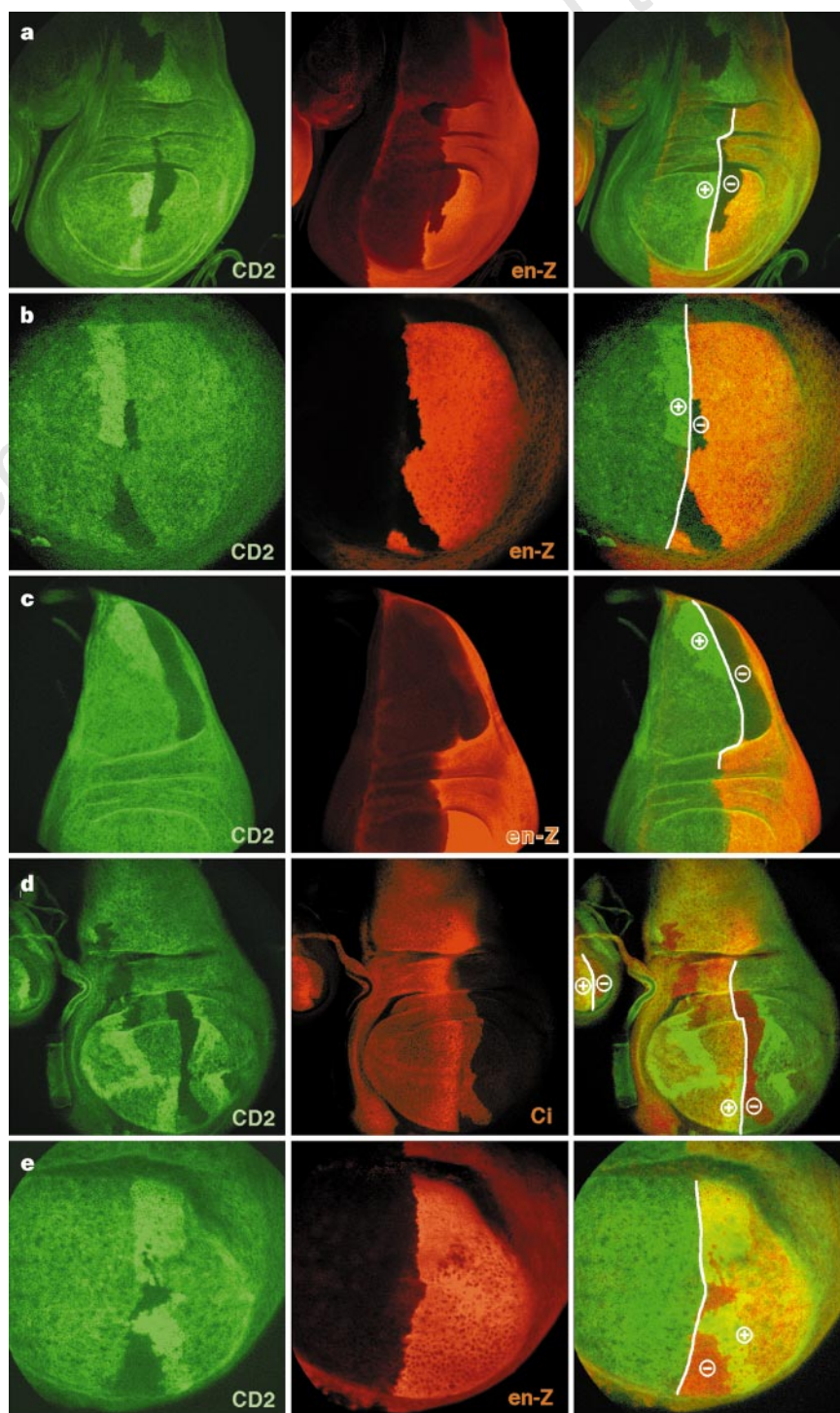
* Present address: Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid, Spain.

smo activity. Conversely, if it is the Hh signal that normally induces a local affinity difference between A and P cells, then a block in Hh signal transduction would cause the affected A cells to sort out from wild-type A cells that do receive the Hh signal, and to intermix instead with P cells on the other side of the compartment boundary.

We used FLP-mediated mitotic recombination¹⁸ to generate 'twin spots' composed of sibling *smo*⁻/*smo*⁻ and *smo*⁺/*smo*⁺ clones in a background of *smo*⁺/*smo*⁺ cells in the wing imaginal disc. As described in the legend to Fig. 1, cells of each genotype could be distinguished by the expression of a marker protein, CD2 (refs 19, 20), and scored for their compartmental provenance by the expression of an *en-lacZ* reporter gene²¹ (expressed exclusively in P cells). Of 497 twin spots analysed, 91 included a *smo*⁻/*smo*⁻ clone

which defined a straight boundary in the normal position of the A/P compartment boundary. Of these 91 twin spots, 34 were contained completely within the P compartment, as all of the cells in both sister clones expressed the *en-lacZ* gene, and two fell entirely within the domain of the normal A compartment. However, in the 55 remaining twin spots, the *smo*⁺/*smo*⁺ clone was located entirely within the normal domain of the A compartment, while its *smo*⁻/*smo*⁻ sister clone appeared to sort entirely within the normal domain of the P compartment (Fig. 1). In these 55 twin spots, all cells in each clone lack *en-lacZ* expression, indicating that they remain of the A-compartment type (Fig. 1a-c). The cells in the *smo*⁻/*smo*⁻ clone formed a straight boundary when juxtaposed with their sister *smo*⁺/*smo*⁺ clone or with neighbouring *smo*⁺/*smo*⁺ A cells,

Figure 1 *Smo*⁻ cells redirect the A/P affinity boundary. Each panel shows a wing imaginal disc double-stained for CD2 expression (marker in green) and either *en-lacZ* (en-Z) reporter gene expression (red in a-c, e) or Cubitus interruptus (Ci) expression (red in d). The merged image is shown to the right. *Smo*⁻/*smo*⁻ clones are marked by the lack of CD2 expression, and their *smo*⁺/*smo*⁺ sister clones can be recognized by the elevated levels of CD2 expression. a-d, Examples of *smo*⁻ clones that originated in the A compartment (as judged by the position of their *smo*⁺/*smo*⁺ sister clone and the absence of *en-lacZ* expression (a-c) or the presence of Ci expression (d)), yet are clearly located in territory normally occupied by P-compartment cells. Such *smo*⁻ clones form straight, smooth boundaries with anteriorly located cells (white line), but form wiggly borders with P-compartment cells. This behaviour of *smo*⁻ cells is observed not only in the wing-blade primordium (a, b, d) but also in the notum (c) and haltere primordia (d, left). *Smo*⁻ clones that originated in the P compartment (e) behave neutrally; they can define the affinity boundary (white line) towards anteriorly located cells, but retain *en-lacZ* expression. Here and in Fig. 3 the discs are oriented anterior to the left and dorsal side up. Discs in b and e are shown at twice the magnification of the other discs. The exact genotypes are given in the Methods section.



but formed a wiggly boundary with neighbouring smo^-/smo^+ cells in the P compartment. The properties of these $smo^+/smo^+ : smo^-/smo^-$ twin spots differ from those of 'neutral' ($smo^+/smo^+ : smo^+/smo^+$) twin spots, both clones of which invariably respected the normal A/P compartment boundary from the same side.

Similar results were also obtained when we analysed $smo^+/smo^+ : smo^-/smo^-$ twin spots in the adult wing^{22,23}. As in the wing imaginal disc, smo^-/smo^- clones generally defined a straight line in the normal position of the A/P boundary from the posterior side, irrespective of whether they originated from A- or P-compartment cells (as scored by the position of their sister clones; Fig. 2c, d). If derived from an A cell, these posteriorly situated smo^- clones invariably generated ectopic A-compartment pattern (Fig. 2b, b). This can best be observed along the wing margin. Unlike anterior cells, posterior cells normally do not differentiate innervated sen-

sory organs with socketed bristles, a difference that has been shown to be controlled by the state of *en* activity^{7,8,24}. We found that smo^- cells autonomously form anterior wing margin structures if they are derived from A cells, even when they are located in the domain normally occupied by P-compartment cells.

Thus, if the reception of the Hh signal is blocked in a clone of A cells by removing Smo activity, these cells appear to intermix with P-compartment cells, suggesting that they have a common affinity with these cells (which are also refractory to Hh). This intermixing occurs despite the different state of En expression in the two cell populations ('off' in the smo^- A cells and 'on' in the P cells). Conversely, cells belonging to these smo^- A clones form a remarkably straight boundary with neighbouring, wild-type A-compartment cells located just anterior to them. It has previously been shown that the absence of Smo function prevents A cells from sequestering as well as transducing Hh, allowing Hh to move across the clone until it is received by smo^+ A-compartment cells on the other side¹⁷. Hence, this straight boundary seems to indicate the formation of an 'affinity boundary' between adjacent smo^+ cells, which receive Hh, and smo^- cells, which cannot, even though En is 'off' in both populations.

To investigate further whether the loss of Hh signal transduction is responsible for the intermixing between smo^- A cells and wild-type P cells, we generated $smo^+/smo^+ : smo^- PKA^-/smo^- PKA^-$ twin spots in which the Hh signal-transduction pathway is constitutively

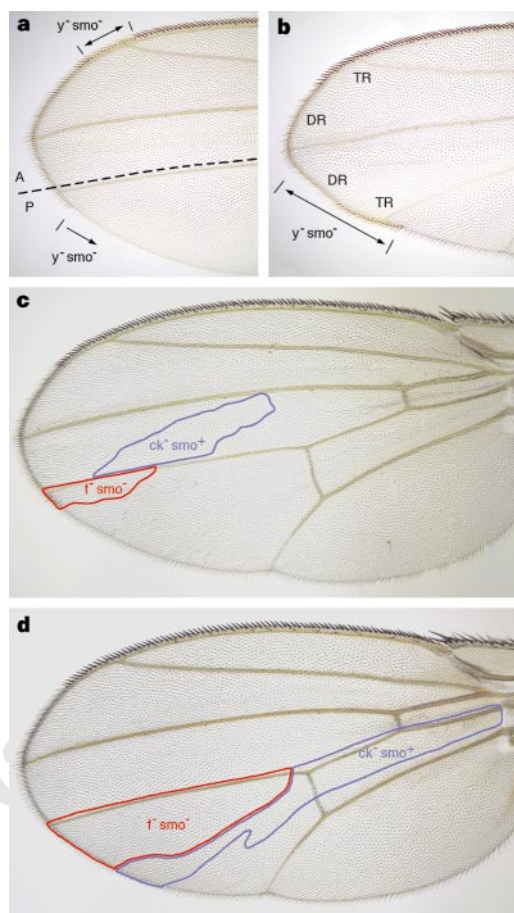


Figure 2 *Smo*⁻ cells that crossed the affinity boundary autonomously form ectopic A pattern. **a, b**, *Smo*⁻ clones marked by the lack of *y*⁺ activity. **a**, Morphologically normal wing with two *smo*⁻ clones located away from the compartment boundary (the position of which is indicated by a broken line just anterior to vein L4). **b**, Clones showing ectopic A pattern are always located posteriorly to, but in the vicinity of, the normal position of the A/P boundary. Ectopic double-row bristles (which have sockets; DR) and triple-row bristles (comprising stout bristles; TR) are formed. Clones at this position can also be phenotypically neutral, if they originate from P-compartment cells (see below). **c, d**, Twin-spot analysis. Different cell markers were used to label the two daughter cells produced after a mitotic recombination event, allowing us to identify and distinguish the resulting sister clones (the exact genotypes are given in the Methods section). *Smo*⁻ clones that abut the A/P boundary from the posterior side, and those that are phenotypically neutral are always accompanied by a sister clone located in the P compartment. However, *smo*⁻ clones that differentiate ectopic anterior structures invariably have their sister clone in the A compartment despite their location in P territory.

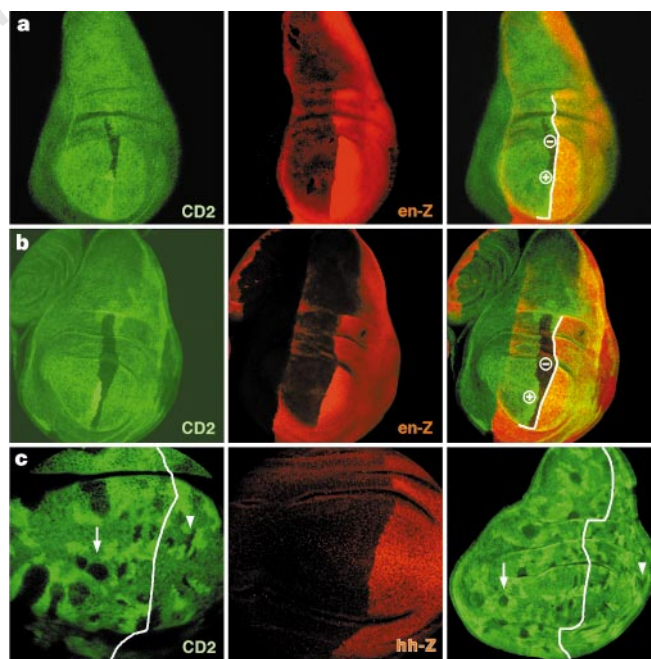


Figure 3 Constitutive activation of the Hh signal-transduction pathway can rescue the abnormal segregation behaviour of *smo*⁻ cells and cause *smo*⁺ A cells to minimize contact with surrounding cells that do not receive the Hh signal. **a, b**, *smo*⁻ *PKA*⁻ double-mutant clones, marked by the lack of CD2 expression (green), form straight boundaries when juxtaposed against P-compartment cells (marked by *en-lacZ* (*en-Z*) expression, red). This contrasts with the behaviour of *smo*⁻ single-mutant cells, which readily mix with P-compartment cells (see Fig. 1). Both clones shown originated from A-compartment cells, as assessed by the location of the sister clone and their state of *en* expression. **c**, *ptc* mutant clones, marked by the lack of CD2 expression (green), are round if located in anterior positions within the A compartment. Compare the shape of anteriorly positioned clones (arrow) with clones located close to, or posterior to, the compartment boundary (arrowhead). The position of the compartment boundary (white line) was assessed by double-staining for the expression of a *hh-lacZ* (*hh-Z*) reporter gene (red). The two images on the left show the same disc as in the image to the right, but at twice the magnification.

activated in the *smo* mutant cells by the simultaneous loss of activity of protein kinase A (PKA)^{17,20,25,26}. Of 43 *smo*⁻ PKA⁻ clones that touched the A/P boundary, 18 formed a straight boundary with neighbouring P-compartment cells and had sister clones in the A compartment (Fig. 3a, b). Five *smo*⁻ PKA⁻ clones with anterior sister clones were not located entirely in the normal domain of the A compartment or in the normal domain of the P compartment. All 20 remaining clones had a posterior sister clone and abutted the A/P boundary from the posterior side. Thus activation of the Hh signal-transduction pathway by elimination of PKA activity changed the segregation behaviour of *smo*⁻ cells, although anterior-compartment affiliation was not in all cases fully restored (possibly because loss of PKA activity does not fully activate the Hh signal-transduction pathway²⁰).

Taken together, our results indicate that the main affinity difference responsible for maintaining the physical segregation between A and P cells is not a heritable, autonomous property under direct control of the selector gene *en*. Rather, it is a result of an inductive interaction that occurs unidirectionally across the common interface between the two cell populations. Even small *smo*⁻ clones induced late in development in the A compartment appear to sort into the P compartment. This indicates that Hh signalling is required throughout the development of the disc to prevent intermixing between A- and P-compartment cells. Our results also suggest that the cell affinities among A-compartment cells are not uniform, as only A-compartment cells in the vicinity of the A/P compartment boundary normally receive high levels of Hh and should therefore differ in affinity from the remaining A-compartment cells, which receive little or no Hh. However, A cells along the compartment boundary that receive Hh mix readily with more anteriorly located cells that are exposed to less Hh. Hence, we infer that only the abrupt apposition of two distinct affinity states prevents cells from intermixing. To reinforce this interpretation we note that clones in which the Hh transduction pathway is constitutively activated by elimination of *patched* (*ptc*) activity²⁷ form smooth boundaries in anterior regions of the A compartment where the surrounding wild-type cells are exposed to little if any Hh (Fig. 3c; see also refs 28, 29). The anterior border of *ptc* mutant clones situated just anterior to the A/P boundary differs according to the allele used¹⁷. It is straight in *ptc*^{S2} clones, apparently because Ptc^{S2} protein on the surface of the mutant cells sequesters Hh¹⁷, preventing it from reaching neighbouring, wild-type cells. But it is wiggly in *ptc* null clones¹⁷, presumably because cells just anterior to the clone are receiving ectopic Hh, which was not sequestered by Ptc protein, and hence have the 'same' affinity as the *ptc*⁻ cells.

The establishment of an A/P compartment boundary can therefore be viewed as a two-step process. First, the selective activation of *en* in a subset of cells creates two cell populations with mutually exclusive properties: Hh-sending, and Hh-receiving, cells. The unidirectional flow of Hh signal from one population to the other causes a local affinity difference in Hh-receiving cells that appears to be predominantly responsible for the subsequent physical segregation of the two cell populations. This same Hh signal also induces the expression of long-range morphogens, such as Dpp, along the boundary. The use of Hh signalling for both purposes may ensure that the position and shape of the morphogen source that organizes both compartments is stably maintained during development. □

Methods

Mutations and markers. All alleles and markers used have been described^{17,18,23}. For most experiments, the *smo*³ allele was used^{14,30}, although other alleles (*smo*¹, *smo*²) were also tested and the same results were obtained. In addition to *en-lacZ* (Figs 1a–c, e, 3a, b), *Ci* (Fig. 1d) and *hh-lacZ* (Fig. 3c), we also used antibodies to En protein to mark P-compartment cells (not shown) and found the same results in all cases.

Generation and analysis of clones. For the imaginal-disc analysis, clones

were generated by Flp-mediated mitotic recombination, subjecting first-instar larvae to a 36 °C heat shock for 30 min. Larvae of the following genotypes were used: *y w hsp70-flp, smo³ FRT39 en^{Xho25}/CD2L.1 FRT39*; *y w hsp70-flp, smo³ DCO^{E95} FRT39 en^{Xho25}/CD2L.1 FRT39*; *y hsp70-flp, FRT42 ptc^{11W}/FRT42 CD2R.1, hh^{P30}/+*. *hsp70-CD2* gene expression was induced by heat-shocking third-instar larvae at 37 °C for 60 min followed by a recovery at 25 °C for 60 min before dissection. Discs were stained using standard immunofluorescence protocols and mouse anti-CD2 (OX34, Serotec), rabbit anti-βGal (Cappel) and anti-Ci (ref. 31) antisera. For adult clones we used the genotype *y w hsp70-flp, smo³ FRT40/P[y⁺]/FRT40*. For the twin-spot analysis in adult wings, X-rays were used to induce somatic recombination. This reduced the clone frequency to ~1 clone per 20 wings, thus avoiding the generation of multiple clones per wing, which can complicate the assignment of sister clones. Larvae of the genotype *f^{36a} hsp70-flp, smo³ FRT40/P[f⁺]/ck FRT40* were used. We scored 1,470 mounted wings and identified 77 *smo*⁻ clones, of which 48 had their sister clone in the A compartment, and 29 in the P compartment. Of the 48 clones with anterior sister clones, 26 were phenotypically neutral and located anterior to vein L3; six were located in the territory between L3 and L4 (and showed phenotypes such as extra vein tissue or a narrower intervein area); and 16 were located in regions normally occupied by P-compartment cells. This latter group invariably exhibited ectopic A pattern (such as anterior wing margin bristles and extra L2 vein) and formed a straight boundary with more anteriorly located cells.

Received 9th July; accepted 22nd August 1997.

- Garcia-Bellido, A., Ripoll, P. & Morata, G. Developmental compartmentalization of the wing disk of *Drosophila*. *Nature* **246**, 251–253 (1973).
- Garcia-Bellido, A., Ripoll, P. & Morata, G. Developmental compartmentalization in the dorsal mesothoracic disc of *Drosophila*. *Dev. Biol.* **48**, 132–147 (1976).
- Morata, G. & Lawrence, P. A. Control of compartment development by the engrailed gene in *Drosophila*. *Nature* **255**, 614–617 (1975).
- Kornberg, T., Siden, L., O'Farrell, P. & Simon, M. The engrailed locus of *Drosophila*: in situ localization of transcripts reveals compartment-specific expression. *Cell* **40**, 45–53 (1985).
- Basler, K. & Struhl, G. Compartment boundaries and the control of *Drosophila* limb pattern by Hedgehog protein. *Nature* **368**, 208–214 (1994).
- Tabata, T. & Kornberg, T. B. Hedgehog is a signaling protein with a key role in patterning *Drosophila* imaginal discs. *Cell* **76**, 89–102 (1994).
- Zecca, M., Basler, K. & Struhl, G. Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the *Drosophila* wing. *Development* **121**, 2265–2278 (1995).
- Tabata, T., Schwartz, C., Gustavson, E., Ali, Z. & Kornberg, T. B. Creating a *Drosophila* wing de novo, the role of engrailed, and the compartment border hypothesis. *Development* **121**, 3359–3369 (1995).
- Capdevila, J. & Guerrero, I. Targeted expression of the signaling molecule decapentaplegic induces pattern duplications and growth alterations in *Drosophila* wings. *EMBO J.* **13**, 4459–4468 (1994).
- Garcia-Bellido, A. Genetic control of wing disc development in *Drosophila*. *Ciba Found. Symp.* **161**–182 (1975).
- Lawrence, P. A. & Struhl, G. Further studies of the engrailed phenotype in *Drosophila*. *EMBO J.* **1**, 827–833 (1982).
- Blair, S. S. Compartments and appendage development in *Drosophila*. *Bioessays* **17**, 299–309 (1995).
- Lawrence, P. A. & Struhl, G. Morphogens, compartments, and pattern: lessons from *Drosophila*. *Cell* **85**, 951–961 (1996).
- Nüsslein-Volhard, C., Wieschaus, E. & Kluding, H. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Wilhelm Roux Arch. Dev. Biol.* **193**, 267–282 (1984).
- Alcedo, J., Ayzenzon, M., Von Ohlen, T., Noll, M. & Hooper, J. E. The *Drosophila* smoothened gene encodes a seven-pass membrane protein, a putative receptor for the hedgehog signal. *Cell* **86**, 221–232 (1996).
- van den Heuvel, M. & Ingham, P. W. *smoothened* encodes a receptor-like serpentine protein required for hedgehog signalling. *Nature* **382**, 547–551 (1996).
- Chen, Y. & Struhl, G. Dual roles for Patched in sequestering and transducing Hedgehog. *Cell* **87**, 553–563 (1996).
- Xu, T. & Rubin, G. M. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117**, 1223–1237 (1993).
- Dunin, B. O. & Brown, N. H. Mammalian CD2 is an effective heterologous marker of the cell surface in *Drosophila*. *Dev. Biol.* **168**, 689–693 (1995).
- Jiang, J. & Struhl, G. Protein kinase A and Hedgehog signaling in *Drosophila* limb development. *Cell* **80**, 563–572 (1995).
- Hama, C., Ali, Z. & Kornberg, T. B. Region-specific recombination and expression are directed by portions of the *Drosophila* engrailed promoter. *Genes Dev.* **4**, 1079–1093 (1990).
- Lawrence, P. A. *Drosophila: a Practical Approach* (eds Roberts, D. B.) 229–242 (IRL, Oxford, 1986).
- Burke, R. & Basler, K. Dpp receptors are autonomously required for cell proliferation in the entire developing *Drosophila* wing. *Development* **122**, 2261–2269 (1996).
- Lawrence, P. A. & Morata, G. Compartments in the wing of *Drosophila*: a study of the engrailed gene. *Dev. Biol.* **50**, 321–337 (1976).
- Li, W., Ohlmeyer, J. T., Lane, M. E. & Kalderon, D. Function of protein kinase A in hedgehog signal transduction and *Drosophila* imaginal disc development. *Cell* **80**, 553–562 (1995).
- Pan, D. & Rubin, G. M. cAMP-dependent protein kinase and hedgehog act antagonistically in regulating decapentaplegic transcription in *Drosophila* imaginal discs. *Cell* **80**, 543–552 (1995).
- Ingham, P. W., Taylor, A. M. & Nakano, Y. Role of the *Drosophila* *patched* gene in positional signalling. *Nature* **353**, 184–187 (1991).
- Phillips, R. G., Roberts, I. J., Ingham, P. W. & Whittle, J. R. The *Drosophila* segment polarity gene *patched* is involved in a position-signalling mechanism in imaginal discs. *Development* **110**, 105–114 (1990).
- Johnson, R. L., Grenier, J. K. & Scott, M. P. Patched overexpression alters wing disc size and pattern: transcriptional and post-transcriptional effects on hedgehog targets. *Development* **121**, 4161–4170 (1995).

30. Lindsley, D. L. & Zimm, G. G. *The Genome of Drosophila melanogaster* (Academic, San Diego, 1992).
 31. Motzny, C. K. & Holmgren, R. The *Drosophila cubitus interruptus* protein and its role in the wingless and hedgehog signal transduction pathways. *Mech. Dev.* 52, 137–150 (1995).

Acknowledgements. We thank J. Nüsslein-Volhard, Y. Chen and G. Struhl for fly stocks; R. Burke, C. Dahmann, N. Méthot and particularly G. Struhl for comments on the manuscript; F. Cifuentes, D. Nellen and M. Zecca for help with the clonal analysis; and J. Modolell for support. I.R. was the recipient of an EMBO fellowship. This work was supported by the Swiss National Science Foundation, the Dirección general de Investigación Científica y Técnica, and the Fundación Ramon Areces.

Correspondence and requests for materials should be addressed to K.B. (e-mail: basler@zool.unizh.ch).

Opposing BMP and EGF signalling pathways converge on the TGF- β family mediator Smad1

Marcus Kretzschmar, Jacqueline Doody & Joan Massagué

Cell Biology and Genetics Program and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

The growth factor TGF- β , bone morphogenetic proteins (BMPs) and related factors regulate cell proliferation, differentiation and apoptosis, controlling the development and maintenance of most tissues^{1,2}. Their signals are transmitted through the phosphorylation of the tumour-suppressor SMAD proteins by receptor protein serine/threonine kinases (RS/TKs)^{3–10}, leading to the nuclear accumulation^{5,9,11,12} and transcriptional activity of SMAD proteins^{6,12,13}. Here we report that Smad1, which mediates BMP signals, is also a target of mitogenic growth-factor signalling through epidermal growth factor and hepatocyte growth factor receptor protein tyrosine kinases (RTKs). Phosphorylation occurs at specific serines within the region linking the inhibitory and effector domains of Smad1, and is catalysed by the Erk family of mitogen-activated protein kinases. In contrast to the BMP-stimulated phosphorylation of Smad1, which affects carboxy-terminal serines and induces nuclear accumulation of Smad1⁶, Erk-mediated phosphorylation specifically inhibits the nuclear accumulation of Smad1. Thus, Smad1 receives opposing regulatory inputs through RTKs and RS/TKs, and it is this balance that determines the level of Smad1 activity in the nucleus, and so possibly the role of Smad1 in the control of cell fate.

Members of the transforming growth factor (TGF)- β family participate in signalling networks that control cell fate in virtually all tissues. Signalling through these networks involves the synergistic or antagonistic interplay of different pathways. One example is the ability of tyrosine kinase receptor activators and BMPs to oppose each others' actions during organ development. Fibroblast growth factor (FGF) opposes the antiproliferative effect of Bmp-2 during limb bud outgrowth¹⁴ and the ability of Bmp-4 to induce interdigital membrane apoptosis during digit formation¹⁵, whereas epidermal growth factor (EGF) can oppose the induction by Bmp-2 of osteogenic differentiation markers¹⁶. Bmp-2 and Bmp-4 can both counteract the induction by FGF of genes essential for tooth development¹⁷. The mechanisms underlying these antagonistic interactions are unknown. However, the discovery^{3,4} of SMAD proteins as mediators of TGF- β family signalling provides a basis to investigate this problem, especially the anti-BMP effects of tyrosine kinase receptor agonists.

SMAD proteins consist of conserved amino- and carboxy-terminal domains (N and C domains, respectively) linked by a more divergent region. The C domain has an effector function that becomes apparent when the isolated domain is tested in transcriptional assays¹² and mesoderm induction assays¹¹. The N domain

interacts physically with the C domain¹⁸, inhibiting its effector activity^{11,12,18}, and contributes to DNA binding¹⁹. Receptor-mediated phosphorylation of serine residues at the end of the C domain (Fig. 1b) relieves it from the inhibitory action of the N domain^{6,9}. Smad1 is activated in this fashion by Bmp-2 and Bmp-4 receptors^{5,6}, and Smad2 and Smad3 are activated by TGF- β and activin receptors^{7,9,20}. Receptor-mediated phosphorylation allows these SMAD proteins to associate with the related Smad4 protein^{6,7,21}, move into the nucleus^{5,6,9,12}, and activate gene responses^{6,7,9,12,13,20}. Smad4 is the product of the tumour-suppressor gene *deleted in pancreatic carcinoma-4* (*DPC4*)²², and is a trimeric protein²³ that is required for signalling by different members of the TGF- β family^{7,24}.

We have previously shown that Smad1 is phosphorylated in proliferating mink lung epithelial cells (R-1B/L17), even in the absence of BMP signals⁶. To investigate the nature of this basal phosphorylation we isolated Smad1 from transiently transfected and ³²P-phosphate-labelled R-1B/L17 cells via a Flag epitope and subjected the protein to proteolytic digestion by various enzymes. Separation by SDS-PAGE revealed ³²P-phosphate-labelled fragments of relative molecular mass 15,000 and 17,000 (*M_r* 15K and 17K) in the digests with trypsin and endopeptidase-LysC, respectively, whereas no fragment larger than 3K could be detected in a chymotryptic digest (Fig. 1a). The predicted digestion pattern of Smad1 allowed us to identify the phosphorylated region as part of the linker region of Smad1 that connects the highly conserved N and C domains (Fig. 1b). This identification was confirmed by microsequencing of the ³²P-phosphate-labelled tryptic fragment purified by high-performance liquid chromatography (HPLC). The C terminus of Smad1, which is phosphorylated directly by the BMP type I receptor at the SSVS sequence⁶ (Fig. 1b), is cleaved by each of these enzymes into a peptide of 2K or less (Fig. 2c).

Smad1 contains predominantly phosphoserine, with some phosphothreonine but not phosphotyrosine⁵ (data not shown). To determine the exact sites of phosphorylation in the linker region, almost all serines and threonines in this region were mutated, either singly or in combination, as substitutions of serine to alanine or threonine to valine (Fig. 1b). Analysis of these mutants in transfected cells demonstrated that none of the mutations had a strong effect on BMP-induced phosphorylation (Fig. 1c, and data not shown). However, the mutant Smad1(4SP/AP), which has mutations in four repeated PXSP motifs, showed a strong reduction in basal phosphorylation (Fig. 1c). Mutation of fewer than four PXSP motifs caused a proportional reduction of phosphorylation (data not shown). When the PXSP motifs in the linker region and the BMP receptor target motif SSVS at the C terminus were mutated simultaneously, phosphorylation of Smad1 was completely lost, both in the presence and absence of stimulation by BMPs (Fig. 1d). This indicates that the PXSP motifs are the linker phosphorylation sites, that they are phosphorylated independently of the C-terminal SSVS motif phosphorylation, and that their phosphorylation is mediated by a kinase other than the BMP receptor.

PXSP motifs are consensus sites for mitogen-activated protein kinases (MAP kinases)²⁵. Because basal phosphorylation of these sites in Smad1 was observed in cells proliferating in the presence of serum, we investigated whether this phosphorylation was induced by specific growth factors that are known to activate MAP kinases. EGF signals through a receptor tyrosine kinase (RTK) and strongly activates the Erk subfamily of MAP kinases, whereas tumour necrosis factor (TNF)- α , an inflammatory cytokine, signals through a different family of receptors and stimulates the stress-activated MAP kinases Jnk and p38 (ref. 26). In cells transfected with Flag-Smad1 and labelled with ³²P-phosphate, EGF induced a rapid increase in Smad1 phosphorylation, whereas TNF- α had no effect (Fig. 2a). Furthermore, mutation of the four PXSP motifs not only reduced the basal phosphorylation of Smad1 in proliferating cells (Fig. 2b, lane 4; see also Fig. 1a, b), but also abolished EGF-induced

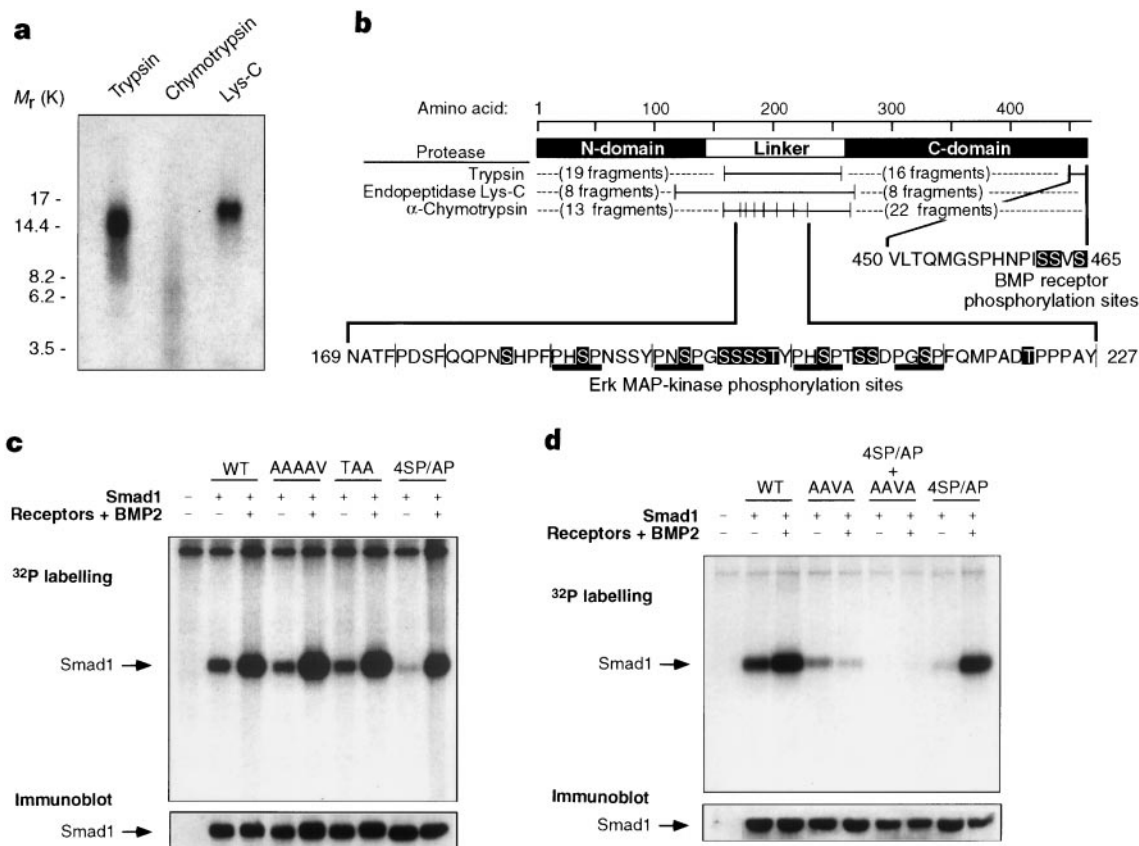


Figure 1 Smad1 is phosphorylated in the linker domain *in vivo*. **a**, Phosphorylated proteolytic fragments correspond by size to the linker region. Fragments were obtained by protease digestion of Smad1 from [32 P]-labelled transfected R-1B/L17 cells. **b**, Predicted proteolytic digestion patterns of Smad1 (ref. 12). The partial sequence of the linker region containing potential phosphorylation sites is shown. Residues mutated singly or in groups are highlighted, and Erk MAP kinase consensus sites are underlined. The sequence of the C-terminal tryptic peptide is also listed with the BMP receptor phosphorylation sites highlighted. **c**,

d, Basal and BMP-induced phosphorylation of the wild-type (WT) Smad1 and selected mutants in transfected R-1B/L17 cells. BMP stimulation was provided by co-transfection of BMP receptors followed by Bmp-2 addition 20 min before cell lysis. Smad1 expression was controlled by anti-Flag immunoblotting. The Smad1 mutants shown contain mutations of S to A or T to V in amino acids 198–202 (AAA), 209–210 (TAA), all four PXSP motifs (serines 187, 195, 206 and 214) (4SP/AP), the C-terminal serines (AAVA), or both the PXSP motifs and the C-terminal serines (4SP/AP + AAVA).

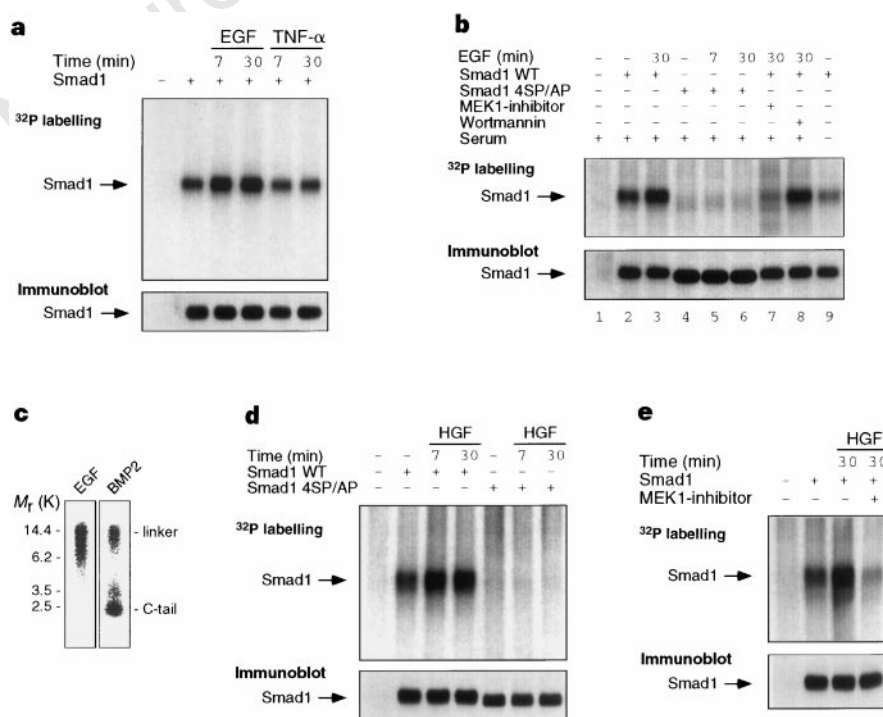


Figure 2 Smad1 is phosphorylated by Erk MAP kinase in response to EGF or HGF. **a**, The effect of EGF and TNF- α on Smad1 phosphorylation in transfected R-1B/L17 cells. [32 P]-labelled cells were incubated with growth factors for the times indicated. **b**, **d**, **e**, Effect of inhibitors and Smad1 PXSP site mutations on Smad1 phosphorylation in response to EGF (**b**) or HGF (**d**, **e**). Inhibitors were added 1 h before growth-factor stimulation. Lane 9 shows the effect of 14 h incubation in serum-free medium. **c**, Trypsin digests of [32 P]-labelled Flag-Smad1 from EGF-treated or Bmp-2-treated cells showing phosphorylation of the linker region but not the C-terminal tail in EGF-treated cells. Phosphorylation of the linker in Bmp-2-treated cells is attributed to the effect of serum factors in the medium.

Smad1 phosphorylation (Fig. 2b, lanes 5 and 6). Addition of EGF did not lead to phosphorylation of the C-terminal sites in Smad1 (Fig. 2c).

Two signalling pathways that mediate the effects of receptor tyrosine kinases are the Ras pathway, which leads to the activation of Erk MAP kinases through Raf and MEK1 (ref. 27), and the phosphatidylinositol-3-OH kinase (PI(3) kinase) pathway, which activates the p70^{S6K} and c-Akt kinases²⁸. Both the basal phosphorylation and the EGF-stimulated phosphorylation of Smad1 were

strongly suppressed by a specific inhibitor of MEK1, whereas wortmannin, an inhibitor of PI(3) kinase, had no effect (Fig. 2b, lanes 7 and 8). To determine whether Smad1 phosphorylation is induced by other growth factors that activate the MAP kinase pathway, we performed similar experiments using hepatocyte growth factor (HGF). HGF is a potent mitogen for epithelial and endothelial cells, signals through the RTK c-Met, and also activates Erk MAP kinases²⁹. Like EGF, HGF rapidly induced Smad1 phosphorylation, and this induction was abolished by mutation of the

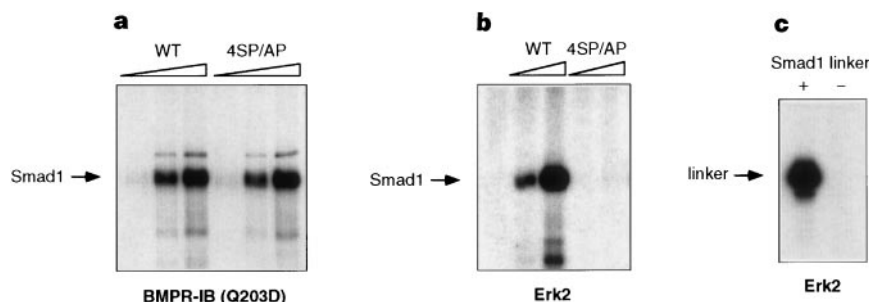


Figure 3 Smad1 is phosphorylated by Erk2 at PXSP sites *in vitro*. **a**, **b**, Purified, recombinant Smad1 proteins (wild-type (WT) or the quadruple PXSP mutant (4SP/AP)) were tested at concentrations in the nanomolar range as substrates of recombinant activated BMP type I receptor kinase (BMPR-IB (Q203D)) or recombinant activated Erk2. **c**, A purified, recombinant Smad1 linker domain was tested as substrate for Erk2.

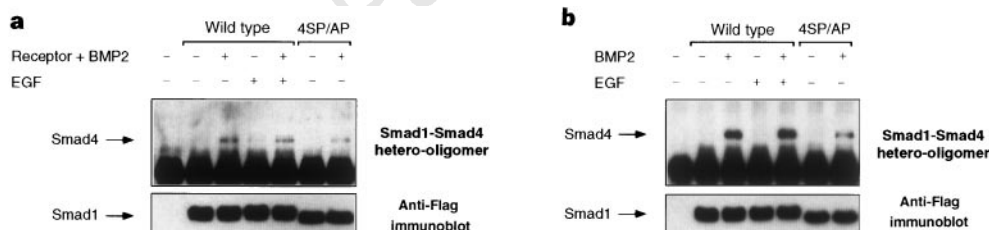
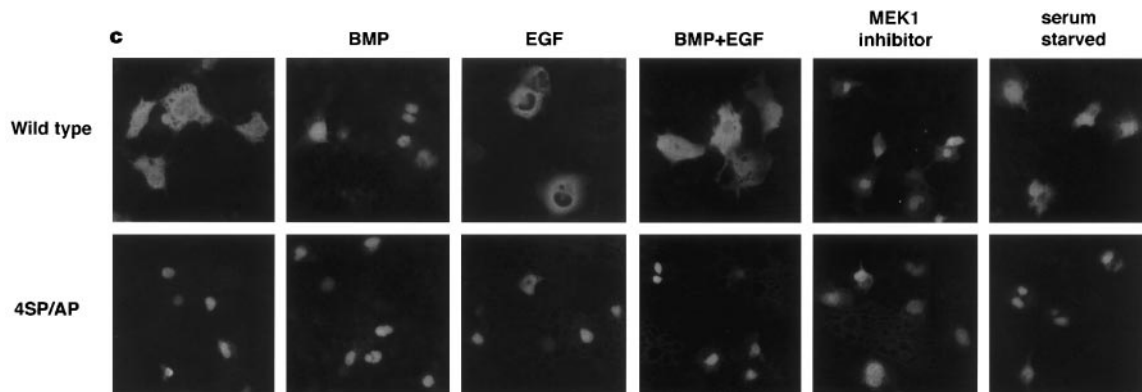
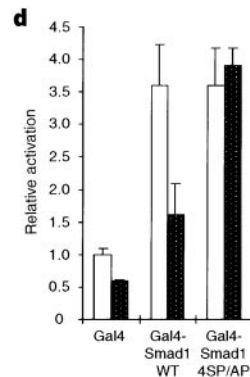


Figure 4 Phosphorylation in the linker region regulates cellular localization and transcriptional function of Smad1 but not association with Smad4. **a**, COS-1 cells were transiently transfected with haemagglutinin-tagged Smad4, and with Flag-tagged Smad1 (wild-type or 4SP/AP mutant) and BMP type I receptor (BMPR-IB (Q203D)) as indicated. Cells were treated with EGF for 15 min before Bmp-2 addition. Then, 1 h after Bmp-2 addition, cells were lysed and subjected to anti-Flag immunoprecipitation followed by anti-haemagglutinin immunoblotting. Smad1 expression levels were controlled by anti-Flag immunoblotting of the same cell lysates. **b**, A similar experiment was performed without transfection of the BMP type I receptor. **c**, COS-1 cells were transiently transfected with Flag-tagged wild-type Smad1 or the mutant Smad1 (4SP/AP). Where indicated, cells were treated with EGF 15 min before Bmp-2 addition. Then, 30 min after Bmp-2 addition, immunostaining was performed using anti-Flag monoclonal antibody and FITC-conjugated secondary antibody. Treatment with MEK1 inhibitor and serum starvation was for 1 h and for 16 h, respectively, before immunostaining. The same slides were counterstained with DAPI to visualize nuclei (data not shown). Several independent experiments were performed with essentially identical results. **d**, R-1B/L17 cells transfected with a Gal4-dependent Luciferase reporter construct and the indicated Gal4 expression constructs were serum starved for 12 h and then treated with or without Bmp-2 in the absence (white bars) or presence (black bars) of EGF. Luciferase activity was measured 18 h after factor addition and plotted as relative activation by Bmp-2. Experiments were performed in triplicate.



four PXSP motifs (Fig. 2c) or the addition of MEK1 inhibitor (Fig. 2d). Taken together, these results provide strong evidence that Smad1 is indeed a target of the Erk subfamily of MAP kinases in the cell. Consistent with this, the basal phosphorylation of Smad1 is reduced upon serum starvation (Fig. 2b, lane 9), a condition that decreases Erk activity. The residual phosphorylation is probably caused by low levels of Erk activity, which remains even after serum starvation.

To provide evidence that Erk kinases can phosphorylate Smad1 directly at the identified sites, we used bacterially expressed, highly purified Smad1 and the mutant Smad1(4SP/AP) as substrates in *in vitro* kinase assays. Both forms of Smad1 were equally well phosphorylated when titrated in kinase reactions with activated BMP type I receptor kinase (Fig. 3a). This is consistent with the observations that the BMP receptor kinase phosphorylates Smad1 at the C-terminal serines⁶ and that the linker mutations do not affect BMP-induced phosphorylation of Smad1 *in vivo* (Fig. 1c, d). Titration of wild-type Smad1 in kinase reactions with activated Erk2 demonstrated that wild-type Smad1 is a substrate for Erk2 at nanomolar concentrations *in vitro* (Fig. 3b). In contrast to the BMP receptor kinase, Erk2 failed to phosphorylate Smad1(4SP/AP) (Fig. 3b). Furthermore, the isolated recombinant linker region of Smad1 was also phosphorylated by Erk2 (Fig. 3c). Thus Erk2 can potentially phosphorylate Smad1 *in vitro* at the same sites that are phosphorylated in response to EGF or HGF *in vivo*.

To understand the functional significance of Smad1 phosphorylation by MAP kinases, we investigated the potential effects on Smad1 activation events. We compared wild-type Smad1 and the mutant Smad1(4SP/AP) for their ability to associate with Smad4/DPC4 in response to BMP signalling^{6,7}. In cells co-transfected with BMP receptors and epitope-tagged Smad1 and Smad4, Smad1 association with Smad4 in response to Bmp-2 was not affected by prior treatment with EGF (Fig. 4a). Even under conditions in which the BMP response was mediated solely by the endogenous BMP receptors, EGF did not significantly alter the association with Smad4 (Fig. 4b). Mutation of the PXSP motifs of Smad1 caused a slight decrease in the BMP-induced association with Smad4 in cells co-transfected with BMP receptors (Fig. 4a), and had a more pronounced effect when the BMP response was mediated solely by endogenous BMP receptors (Fig. 4b). This reduction in BMP-induced association with Smad4 might reflect a decreased access of Smad1(4SP/AP) to the receptor as a result of a preferential localization of this mutant in the nucleus (see below).

In contrast to the relatively weak effects on the association between Smad1 and Smad4, both addition of EGF and the mutations of the PXSP site had striking effects on the subcellular localization of Smad1. Immunofluorescence of Smad1-transfected cells revealed a distribution throughout the cell under basal conditions and a predominantly nuclear localization after Bmp-2 stimulation (Fig. 4c), as previously reported^{5,6,12}. The mutant Smad1(4SP/AP) was predominantly nuclear even in the absence of BMP stimulation (Fig. 4c). Treating the cells with EGF caused exclusion of wild-type Smad1 from the nucleus and inhibited BMP-induced nuclear accumulation of wild-type Smad1 (Fig. 4c), but had no effect on the mutant Smad1 (Fig. 4c). Furthermore, addition of the MEK1 inhibitor or serum starvation affected wild-type Smad1 in an opposite way to EGF treatment, leading to nuclear accumulation even in the absence of treatment with BMP (Fig. 4c). MEK1 inhibitor or serum starvation had no effect on the mutant Smad1 (Fig. 4c). These results suggest that Erk-mediated phosphorylation of the linker region inhibits nuclear accumulation of Smad1 and is capable of preventing BMP-induced nuclear accumulation without interfering with formation of the Smad1–Smad4 complex. By using Smad4-defective cells we confirmed that Smad1 association with Smad4 and Smad1 nuclear translocation are independent events (F. Liu, C. Poupponnot and J.M., unpublished data). Because activated Erk kinases are distributed through the cytoplasm and the nucleus³⁰,

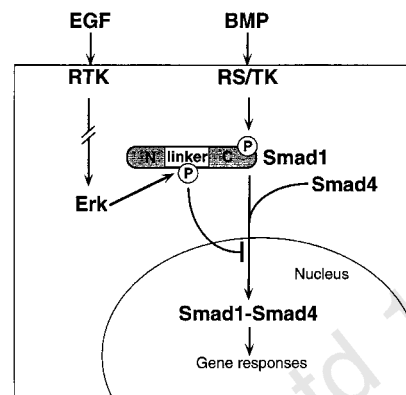


Figure 5 Schematic representation of the regulation of Smad proteins by opposing signalling inputs mediated by receptor tyrosine kinases (RTKs) and receptor serine/threonine kinases (RS/TKs).

phosphorylation of Smad1 by Erk could occur in either compartment and regulate Smad1 recognition by either nuclear import or export mechanisms.

To assess the functional consequences of Smad1 regulation by Erk MAP kinases, we sought to analyse the effects of EGF treatment on the transcriptional activity of Smad1 in response to BMP. Because a Smad1-dependent early response reporter gene for BMP in mammalian cells has not yet been described, we resorted to the previously used Gal4-dependent reporter assay^{6,12}. Cells expressing a fusion of the Gal4 DNA-binding domain and Smad1 (Gal4–Smad1) showed a BMP-inducible increase in Luciferase activity when co-transfected with a Gal4-dependent Luciferase reporter construct (Fig. 4d). BMP had no effect on the Gal4 DNA-binding domain alone (Fig. 4d). The Gal4–Smad1(4SP/AP) mutant construct was induced by BMP to the same extent as Gal4–Smad1 (Fig. 4d), consistent with the observation that the mutant Smad1 is phosphorylated to a similar extent to wild-type Smad1 in response to BMP (Fig. 1c, d). Most importantly, simultaneous treatment of cells with EGF strongly reduced the BMP-induced transcriptional activity of Gal4–Smad1, whereas EGF had no effect on the activity of Gal4–Smad1(4SP/AP) (Fig. 4d). This suggests that Erk-mediated phosphorylation of Smad1 inhibits BMP-induced nuclear accumulation, and consequently inhibits nuclear functions of Smad1, such as transcriptional activity.

Our data show that Smad1, a mediator of BMP receptor signals, is a target of growth-factor signalling through the Erk MAP kinase pathway (Fig. 5). Erk phosphorylates Smad1 at the linker region preventing nuclear accumulation of Smad1 under basal conditions and inhibiting nuclear accumulation and the transcriptional activity of Smad1 in response to BMP. Erk activators should therefore oppose the functions of BMP that depend on nuclear accumulation and transcriptional activity of Smad1. The differential regulation of Smad1 through RTKs and RS/TKs provides a mechanism for antagonistic actions of mitogenic factors and BMPs. This mechanism may underlie, at least in part, the opposing effects of these factors during vertebrate development^{14–17}. Smad2 and Smad3, which are TGF- β /activin mediators, contain similar potential phosphorylation sites in their linker region, and may therefore also be targets of proline-directed kinases (M.K. and J.M., unpublished observations). Therefore SMAD regulation by MAP kinases might be a general phenomenon controlling the signalling activity of the TGF- β family. □

Methods

Transfection, metabolic labelling and immunofluorescence assays. All of these procedures were performed essentially as described⁶. In brief, R-1B/L17 cells were transiently co-transfected with Flag-tagged Smad1 (ref. 12) and the BMP receptors BMPR-IB and BMPR-II^{12,31} as indicated. All constructs were in pCMV5. Smad1 mutant constructs were obtained by standard *in vitro* muta-

genesis procedures. Cells were metabolically labelled 3 days after transfection with [32 P] phosphate for 3 h, treated with the indicated factors (5 nM Bmp-2, Genetics Institute; 18 nM EGF, R&D Systems; 1.1 nM HGF, R&D Systems) and lysed. Where indicated, cells were treated with the MEK1 inhibitor PD98059 (100 μ M, New England Biolabs) or the PI(3) kinase inhibitor wortmannin (0.1 μ M, Calbiochem) for 1 h before addition of growth factor. Serum starvation was for 14 h before metabolic labelling. After cell lysis, Flag-Smad1 was precipitated with monoclonal anti-Flag antibody (M2, Kodak Scientific) and proteins were resolved by SDS-PAGE and visualized by autoradiography. Parallel cultures were treated equivalently, lysed and subjected to western immunoblotting with anti-Flag antibody M2 and by chemiluminescence (ECL, Amersham).

Kinase assays. Smad1 and the Smad1 linker domain (amino acids 146–264) were subcloned into a pET expression vector (Novagen) encoding an N-terminal hexahistidine tag. Bacterial expression and purification of recombinant proteins were performed as described⁶. Smad1 proteins were preincubated (30 min at 4°C) with recombinant, activated BMPR-IB(Q203D) cytoplasmic domain (amino acids 150–502)⁶ or with recombinant, activated Erk2 MAP kinase (New England Biolabs) in a buffer containing 50 mM Tris-HCl, pH 7.3, 100 mM NaCl, 10 mM MnCl₂, 10% (v/v) glycerol, 5 mM dithiothreitol, and 0.05% (v/v) Triton X-100. Upon [γ - 32 P]ATP addition, reactions were performed at 28°C for 20 min. Reactions were stopped by addition of a buffer containing 6 M guanidinium-HCl, and Smad1 was recovered with Ni-NTA agarose.

Phosphopeptide analysis. Smad1 protein was immunoprecipitated from transfected, [32 P]phosphate-labelled R-1B/L17 cells, separated by SDS-PAGE, and transferred to nitrocellulose. Membrane pieces containing Smad1 were incubated with trypsin, chymotrypsin and endopeptidase Lys-C (Promega, Boehringer-Mannheim, and Worthington, respectively) in 50 mM ammonium bicarbonate at 37°C for 3 h. Digests were resolved on 16.5% Tris-trycine gels³² and visualized by autoradiography. M_r markers were 14 C-methylated peptide fragments (Sigma).

SMAD association assay. In SMAD association experiments⁶, COS-1 cells co-transfected with haemagglutinin-tagged Smad4, Flag-tagged Smad1 and BMPR-IB(Q203D) were treated with EGF (18 nM) for 15 min before Bmp-2 addition. Cells were lysed⁶ and portions used in anti-Flag immunoblotting assays to control the Smad1 expression level. The remainder of the cell lysates were immunoprecipitated with anti-Flag antibody. Immunoprecipitates were washed and subjected to anti-haemagglutinin immunoblotting.

Luciferase assays. Luciferase assays were performed essentially as described⁷. In brief, R-1B/L17 cells were transiently cotransfected with a Gal4-luciferase reporter construct³³ (2 μ g) and the indicated Gal4 or Gal4-Smad1 expression constructs (0.5 μ g). Cells were serum starved for 12 h before treatment with Bmp-2 (5 nM) and/or EGF (18 nM), and Luciferase assays were performed 18 h after addition of growth factor.

Received 2 June; accepted 18 August 1997.

1. Roberts, A. B., Sporn, M. B. In *Peptide Growth Factors and their Receptors* (eds Sporn, M. B. & Roberts, A. B.) 419–472 (Springer, Heidelberg, 1990).
2. Hogan, B. L. M. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev.* **10**, 1580–1594 (1996).
3. Eppert, K. *et al.* MADR2 maps to 18q21 and encodes a TGF β -regulated MAD-related protein that is functionally mutated in colorectal carcinoma. *Cell* **86**, 543–552 (1996).
4. Graff, J. M., Bansal, A. & Melton, D. A. *Xenopus* Mad proteins transduce distinct subsets of signals for the TGF β superfamily. *Cell* **85**, 479–487 (1996).
5. Hoodless, P. A. *et al.* MADR1, a MAD-related protein that functions in BMP2 signalling pathways. *Cell* **85**, 489–500 (1996).
6. Kretschmar, M., Liu, F., Hata, A., Doody, J. & Massagué, J. The TGF β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev.* **11**, 984–995 (1997).
7. Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massagué, J. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**, 832–836 (1996).
8. Lechleider, R. J., de Caestecker, M. P., Dehejia, A., Polymeropoulos, M. H. & Roberts, A. B. Serine phosphorylation, chromosomal localization and TGF β signal transduction by human *bsp-1*. *J. Biol. Chem.* **271**, 17617–17620 (1996).
9. Macias-Silva, M. *et al.* MADR2 is a substrate of the TGF β receptor and phosphorylation is required for nuclear accumulation and signaling. *Cell* **87**, 1215–1224 (1996).
10. Yingling, J. M. *et al.* Mammalian Dwarfs are phosphorylated in response to TGF- β and are implicated in control of cell growth. *Proc. Natl Acad. Sci. USA* **93**, 8940–8944 (1996).
11. Baker, J. & Harland, R. M. A novel mesoderm inducer, mMad-2, functions in the activin signal transduction pathway. *Genes Dev.* **10**, 1880–1889 (1996).
12. Liu, F. *et al.* A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**, 620–623 (1996).
13. Chen, X., Rubock, M. J. & Whitman, M. A transcriptional partner of MAD proteins in TGF- β signalling. *Nature* **383**, 691–696 (1996).
14. Niswander, L. & Martin, G. R. FGF-4 and BMP-2 have opposite effects on limb growth. *Nature* **361**, 68–71 (1993).

15. Ganan, Y., Macias, D., Duterte-Coquillaud, M., Ros, M. A. & Hurler, J. M. Role of TGF β s and BMPs as signals controlling the position of the digits and the areas of interdigital cell death in the developing chick limb autopod. *Development* **122**, 2349–2357 (1996).
16. Bernier, S. M. & Goltzman, D. Effect of protein and steroidal osteotropic agents on differentiation and epidermal growth factor-mediated growth of the CFK1 osseous cell line. *J. Cell Physiol.* **152**, 317–327 (1992).
17. Neubüser, A., Peters, H., Balling, R. & Martin, G. R. Antagonistic interactions between FGF and BMP signalling pathways: a mechanism for positioning the sites of tooth formation. *Cell* **90**, 247–255 (1997).
18. Hata, A., Lo, R. S., Wotton, D., Lagna, G. & Massagué, J. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature* **388**, 82–87 (1997).
19. Kim, J., Johnson, K., Chen, H. J., Carroll, S. & Laughon, A. *Drosophila* Mad binds to DNA and directly mediates activation of *vestigial* by Decapentaplegic. *Nature* **388**, 304–308 (1997).
20. Zhang, Y., Feng, X.-H., Wu, R.-Y. & Derynck, R. Receptor-associated Mad homologues synergize as effectors of the TGF- β response. *Nature* **383**, 168–172 (1996).
21. Wu, R.-Y., Zhang, Y., Feng, X.-H. & Derynck, R. Heteromeric and homomeric interactions correlate with signalling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol. Cell. Biol.* **17**, 2521–2528 (1997).
22. Hahn, S. A. *et al.* DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* **271**, 350–353 (1996).
23. Shi, Y., Hata, A., Lo, R. S., Massagué, J. & Pavletich, N. P. A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature* **388**, 87–93 (1997).
24. Zhang, Y., Musci, T. & Derynck, R. The tumor suppressor Smad4/DPC4 as a central mediator of Smad function. *Curr. Biol.* **7**, 270–276 (1997).
25. Hill, C. S. & Treisman, R. Transcriptional regulation by extracellular signals: mechanisms and specificity. *Cell* **80**, 199–211 (1995).
26. Vandenabeele, P., Declercq, W., Beyaert, R. & Fiers, W. Two tumour necrosis factor receptors: structure and function. *Trends Cell Biol.* **5**, 392–399 (1995).
27. Marshall, C. J. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80**, 179–185 (1995).
28. Franke, T. F., Kaplan, D. R. & Cantley, L. C. PI3K: downstream AKTion blocks apoptosis. *Cell* **88**, 435–437 (1997).
29. Besser, D. *et al.* Regulation of the urokinase-type plasminogen activator gene by the oncogene Tpr-Met involves GRB2. *Oncogene* **14**, 705–711 (1997).
30. Blenis, J. Signal transduction via the MAP kinases: proceed at your own RSK. *Proc. Natl Acad. Sci. USA* **90**, 5889–5892 (1993).
31. ten Dijke, P. *et al.* Identification of type I receptors for Osteogenic Protein-1 and Bone Morphogenetic Protein-4. *J. Biol. Chem.* **269**, 16985–16988 (1994).
32. Schagger, H. & von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379 (1987).
33. Gu, W., Shi, X.-L. & Roeder, R. G. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819–823 (1997).
34. Massagué, J., Hata, A. & Liu, F. TGF- β signalling through the Smad pathway. *Trends Cell Biol.* **7**, 187–192 (1997).

Acknowledgements. We thank the Genetics Institute for providing BMP; Memorial Sloan-Kettering Cancer Center for use of the Molecular Cytology Core Facility; and I. Timokhina for discussions. This work was supported by NIH grants to J.M. and to Memorial Sloan-Kettering Cancer Center. M.K. is a recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. J.M. is an investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to J.M. (e-mail: j-massague@ski.mskcc.org).

Smad6 inhibits signalling by the TGF- β superfamily

Takeshi Imamura*, Masao Takase*, Ayako Nishihara, Eiichi Oeda, Jun-ichi Hanai, Masahiro Kawabata & Kohei Miyazono

Department of Biochemistry, The Cancer Institute, Tokyo, Japanese Foundation for Cancer Research, and Research for the Future Program, Japan Society for the Promotion of Science, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan

* These authors contributed equally to this work

SMAD proteins¹ have been identified as signalling mediators of the TGF- β superfamily, which is involved in a range of biological activities including cell growth, morphogenesis, development and immune responses^{2,3}. Smad1, Smad2, Smad3 and Smad5 are ligand-specific: Smad1 and Smad5 transduce signals from bone morphogenetic proteins^{4–7}, and Smad2 and Smad3 mediate signalling by TGF- β and activin^{8,9}, whereas Smad4 acts as a common signalling component¹⁰. For example, Smad2 is phosphorylated by the TGF- β type I receptor upon ligand binding, forms a heteromer with Smad4, and then translocates into the nucleus where it activates transcription^{10,11}. Here we report the isolation of Smad6 in the mouse. Smad6 is quite different in structure from the other SMAD proteins, and forms stable associations with type I receptors. Smad6 interferes with the phosphorylation of Smad2 and the

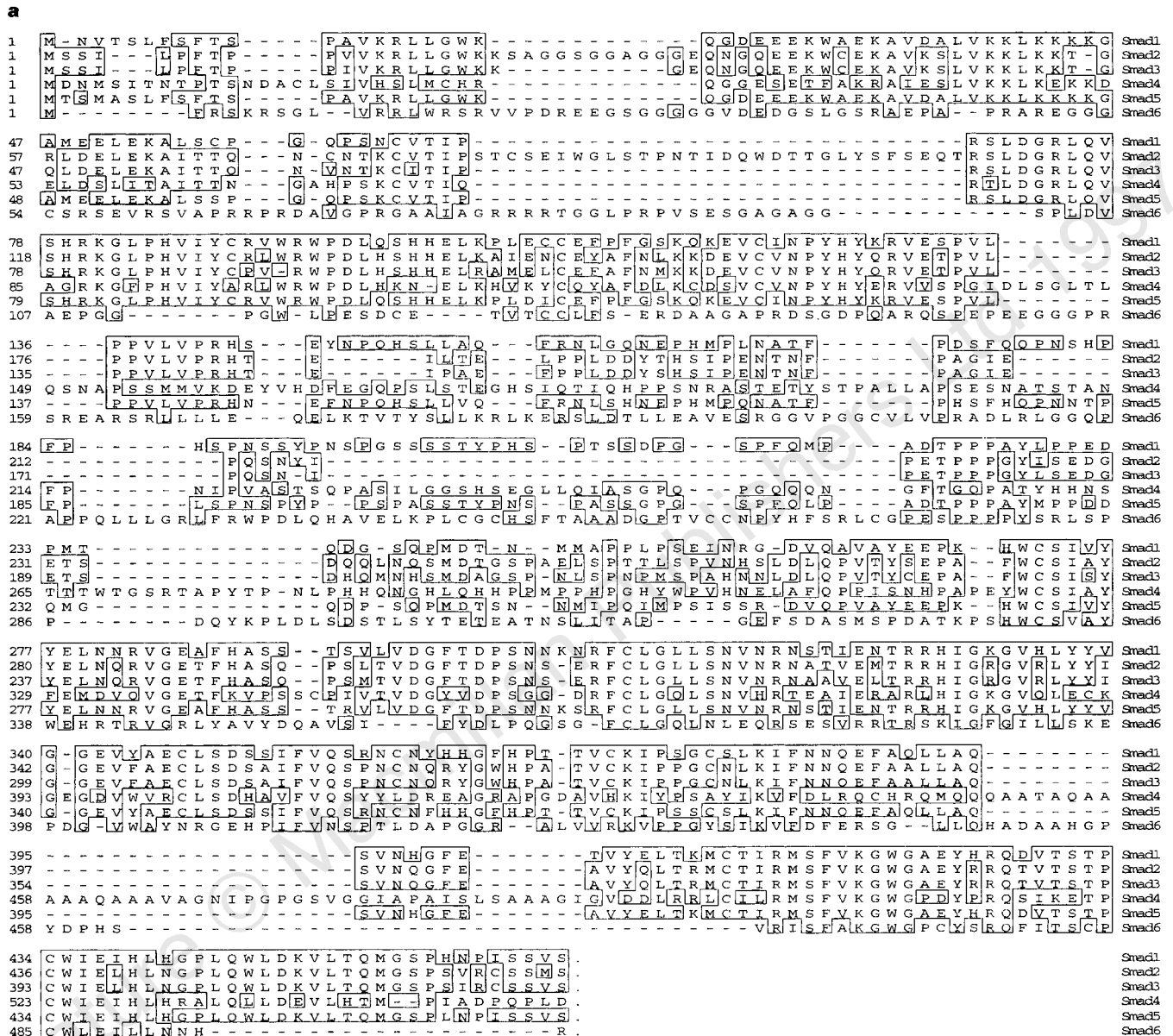


Figure 1 Protein sequence alignment and tissue distribution of mouse Smad6.

a, The deduced amino-acid sequence of Smad6 is aligned with other mammalian SMAD proteins. Identical residues are boxed. The N-terminal two-thirds of mouse Smad6 shows little similarity with the other SMADs. **b**, Human (left) and mouse (right) tissue blots were hybridized with the human EST clone 429356 as a probe. A transcript of 3.0 kb is detected, with the highest expression in lung in both blots.

subsequent heteromerization with Smad4, but does not inhibit the activity of Smad3. Smad6 also inhibits the phosphorylation of Smad1 that is induced by the bone morphogenetic protein type IB receptor. These data indicate that signals of the TGF- β superfamily are regulated both positively and negatively by members of the SMAD family.

We have identified the complete coding sequence of murine Smad6 by screening a mouse lung cDNA library with an expressed-sequence tag (EST) clone (429356) as a probe. Partial sequence information and chromosome localization of the human homologue of this gene has been reported¹² (accession no. U59914). Smad6 is a protein of 495 amino acids with a predicted relative molecular mass of 53,700 (M_r 53.7K) (Fig. 1a). All of the SMAD proteins, except the human homologue, including those of *Drosophila* and *Caenorhabditis elegans*, have conserved amino- and carboxy-terminal regions (MH1 and MH2, respectively) separated by a proline-rich linker region of variable length and sequence, although Smad4 has a unique insert in its MH2 region¹. The C-terminal one-third of Smad6 shares the conserved sequence with the MH2 regions of the other SMAD proteins, but its N-terminal region shows a striking difference from the conserved MH1 sequence (Fig. 1a), suggesting that it has a different function. Northern blot analysis of various human and mouse tissues revealed relatively ubiquitous expression of the mRNA species of 3.0 kilobases (kb), with lung having the highest expression (Fig. 1b).

Members of the transforming growth factor (TGF)- β superfamily exert their diverse effects by binding to two types of receptor with serine/threonine kinase activity³. The ligand first binds to the type II receptor, which consequently activates the type I receptor by direct phosphorylation. The activated type I receptor then phosphorylates ligand-specific SMAD proteins, such as Smad1, Smad2 and Smad3 (refs 1, 9–11). The association of Smad2 with the TGF- β type I receptor (T β R-I) requires activation of T β R-I by the type II receptor (T β R-II)¹¹. Smad2, however, interacts with T β R-I only transiently under physiological conditions, as Smad2 is released from T β R-I after phosphorylation by the receptor. The interaction of Smad2 with T β R-I has thus been observed only when the kinase-defective

form of T β R-I is used¹¹. It should be noted that Smad4 does not associate with the receptors⁹.

We examined the interaction of Smad6 with the type I receptors in affinity crosslinking assays. Smad6 bound to the TGF- β receptor complexes, as indicated by the coprecipitation of the receptor complexes with Smad6 (Fig. 2a). As with Smad2, the binding of Smad6 to T β R-I required the kinase activity of T β R-II (Fig. 2a), although Smad6 stably bound wild-type T β R-I. Similar results were obtained with the activin type IB receptor (ActR-IB) (Fig. 2b) and the bone morphogenetic protein (BMP) type IB receptor (BMPR-IB) (Fig. 2c), for which Smad6 bound to both wild-type and kinase-defective type I receptors. These results suggest that Smad6 binds to the type I receptors in a ligand-dependent manner but has a function different from that of the other SMAD proteins.

We investigated the functional interaction of Smad6 with other SMAD proteins. Smad2 is phosphorylated at its C-terminal end by activated T β R-I (ref. 11). The phosphorylation is essential for the downstream signalling events that culminate in transcriptional activation of the target genes, as disruption of the phosphorylation sites abolished responses induced by TGF- β ¹¹. We then examined the effect of Smad6 on the phosphorylation of Smad2 (Fig. 3b). Phosphorylation of Smad2 induced by constitutively active T β R-I was suppressed by Smad6 (39% reduction, as normalized for the ³²S-labelled band), whereas Smad2 did not affect the constitutive phosphorylation of Smad6 (Fig. 3a). Smad3 and Smad2 have 91% identity in their amino-acid sequence and have been shown to mediate TGF- β signals^{8,9}, although we know of no functional differences between the two molecules. Smad6 enhanced receptor-induced phosphorylation of Smad3 (Fig. 3c), suggesting that Smad6 has different effects on these closely related molecules. We then investigated the effect of Smad6 on Smad1 phosphorylation (Fig. 3d). Smad1 was phosphorylated by both the constitutively active BMPR-IA and BMPR-IB. Smad6 efficiently inhibited phosphorylation induced by the BMPR-IA (60% reduction) but not by BMPR-IB. These results suggest that Smad6 acts as an inhibitor towards certain members of the SMAD family.

Smad2 forms a heteromer with Smad4 upon phosphorylation by

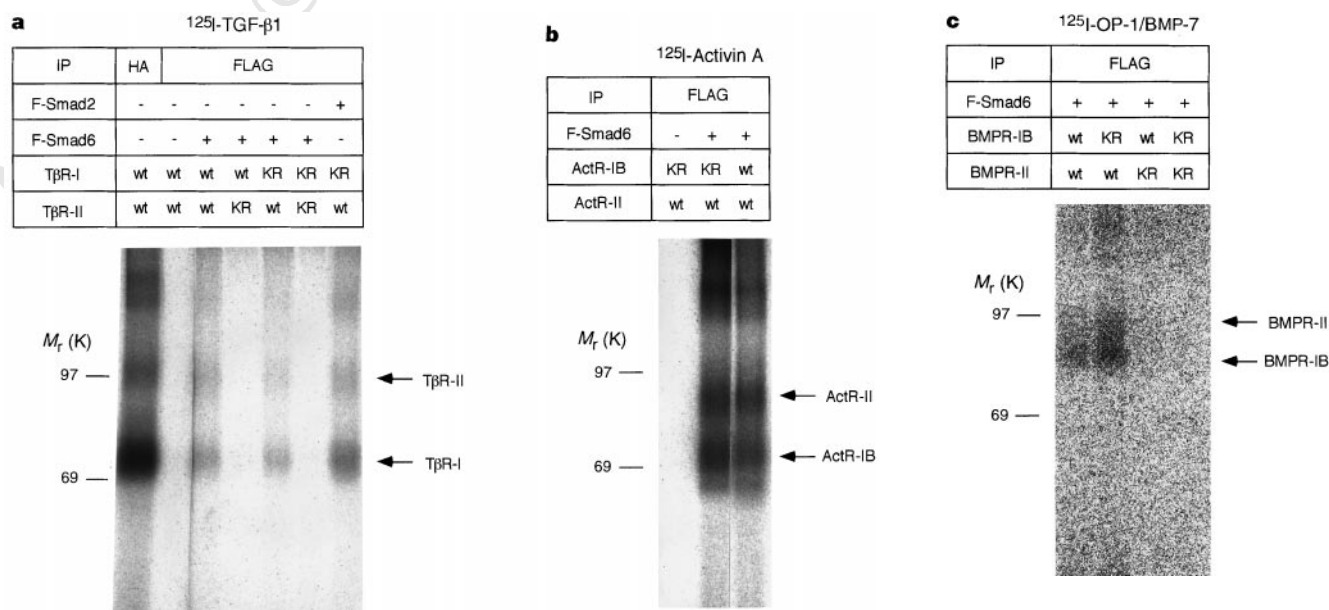


Figure 2 Binding of Smad6 to the type I receptors. **a**, COS-7 cells were transfected with Flag-tagged Smad6 (F-Smad6) or Flag-tagged Smad2 (F-Smad2) in combination with the wild-type (wt) or kinase-defective (KR) haemagglutinin (HA)-tagged T β R-I and hexahistidine-tagged T β R-II. Cells were affinity-labelled with ¹²⁵I-TGF- β 1 and lysates were immunoprecipitated (IP) with anti-HA antibody or anti-FLAG M2 antibody. Immune complexes were subjected to SDS-PAGE

and autoradiography. Smad6 bound to both wild-type and kinase-defective T β R-I, depending on the kinase activity of T β R-II. In contrast, Smad2 bound to kinase-defective T β R-I but not the wild-type T β R-I (data not shown). **b**, A similar experiment was done with ¹²⁵I-activin A. Smad6 associated with both wild-type and kinase-defective ActR-IB. **c**, An experiment using ¹²⁵I-OP-1/Bmp-7 was performed. Smad6 bound to BMPR-IB only when activated by BMPR-II.

Figure 3 Effect of Smad6 on the phosphorylation of Smad2, Smad3 or Smad1 by the constitutively active type I receptors. **a**, COS-7 cells were transiently transfected with constitutively active (TD) T β R-I, Flag-Smad2 (F-Smad2), and/or Myc-Smad6 (M-Smad6). Cells were labelled with [32 P]orthophosphate and lysates subjected to immunoprecipitation (IP) with anti-Myc antibody. Phosphorylated Smad6 was detected by SDS-PAGE and autoradiography. Doublet bands of phosphorylated Smad6 were detected. Neither T β R-I (TD) or Flag-Smad2 affected the phosphorylation. **b**, Cell lysates were immunoprecipitated with anti-Flag antibody to detect Smad2 phosphorylation. Expression of Smad6 suppressed phosphorylation of Smad2 induced by T β R-I (TD). Expression levels of Smad6 (**a**) and Smad2 and T β R-I (TD) (**b**) were monitored by labelling the cells with [35 S]methionine/cysteine. **c**, A similar experiment was done with Smad3. Smad6 did not suppress receptor-induced phosphorylation of Smad3. **d**, Smad1 was phosphorylated by constitutively active (QD) BMPR-1A as well as by BMPR-1B. Smad6 inhibited phosphorylation of Smad1 that was induced by BMPR-1B, but not by BMPR-1A.

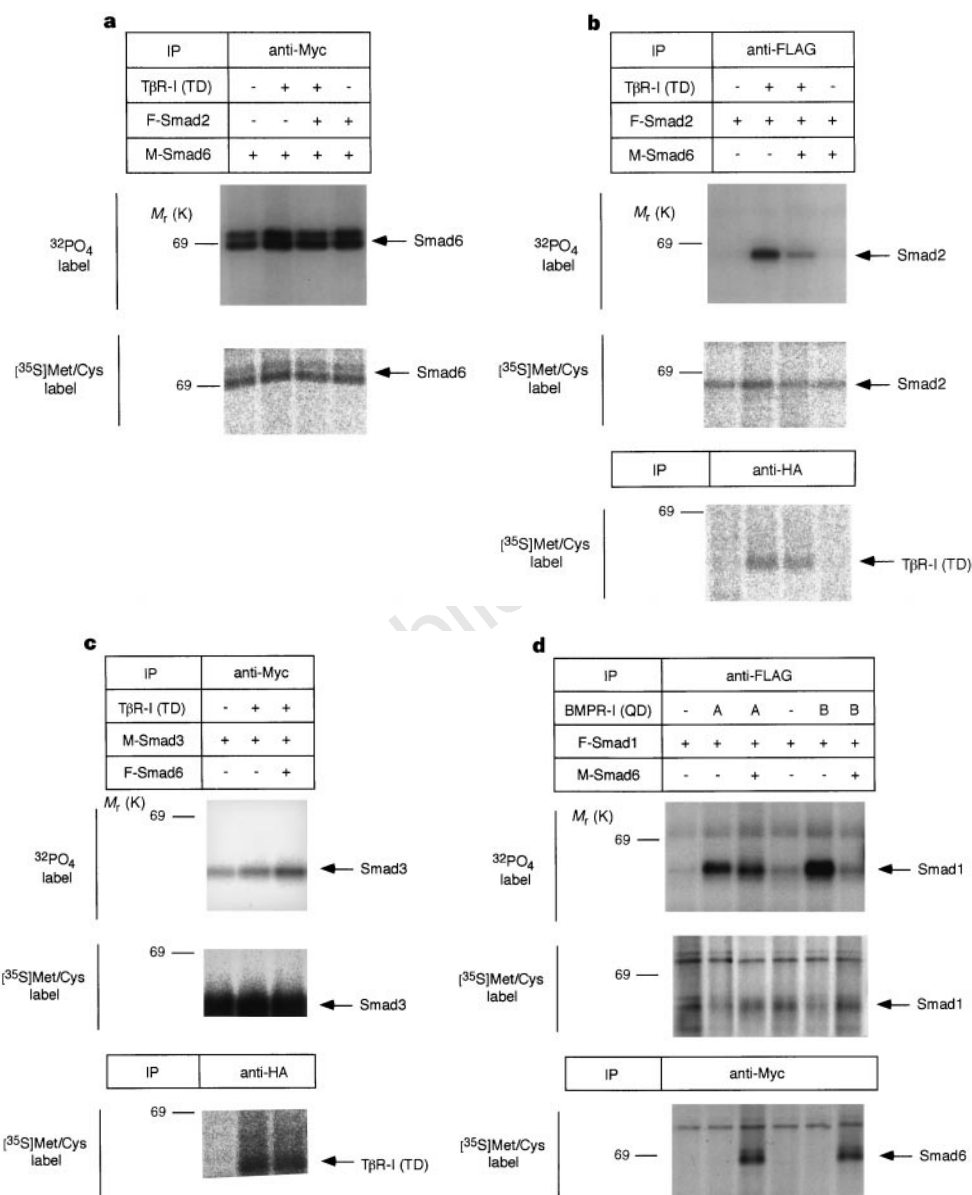
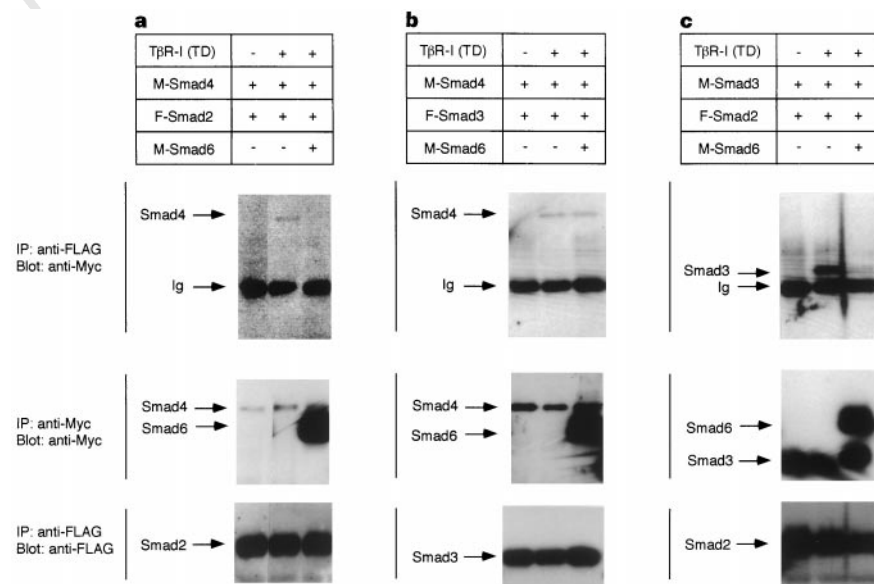


Figure 4 Effect of Smad6 on the heteromerization of Smad2, Smad3 and Smad4. **a**, COS-7 cells were transfected with the indicated combination of plasmids and subjected to immunoprecipitation (IP) followed by western blot detection. Myc-Smad4 (M-Smad4) and Flag-Smad2 (F-Smad2) interacted in the presence of T β R-I (TD), but this was abolished by the expression of Myc-Smad6 (top). Expression levels of Smad2 (bottom), Smad4 (middle), and Smad6 (middle) were monitored. Note that Smad6 did not interact with Smad2 under these conditions (top). **b**, Smad6 did not affect receptor-induced association of Smad3 and Smad4. Smad4 coprecipitated with Smad3 both with and without Smad6 (top). **c**, T β R-I (TD) induced association of Smad2 and Smad3, but this was abolished by Smad6. Smad3 coprecipitation with Smad2 disappeared in the presence of Smad6 (top). Ig, immunoglobulin.



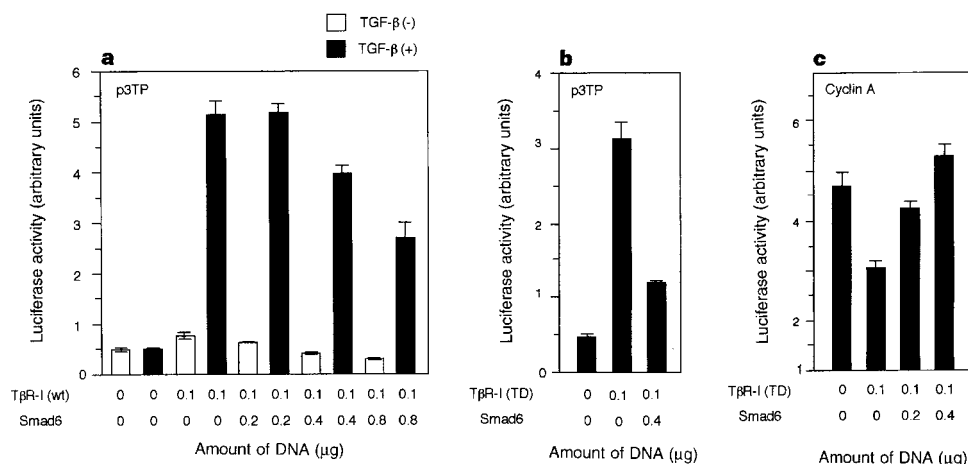


Figure 5 Effect of Smad6 on the transcriptional responses of TGF- β . **a**, Mink R mutant cells deficient in T β R-I were transfected with p3TP-Lux reporter, T β R-I and increasing amounts of Smad6 DNA. Cells were treated with (black bars) or without (white bars) 5 ng ml⁻¹ TGF- β for 24 h. Smad6 inhibited luciferase activity induced by TGF- β in a dose-dependent manner. **b**, Smad6 also inhibited the activity of T β R-I (TD). **c**, Cyclin A luciferase reporter, pCAL2, was used to examine the effect of Smad6 on TGF- β signalling. T β R-I downregulated cyclin A luciferase activity, which was blocked by Smad6.

T β R-I (ref. 10). It was recently shown that TGF- β also induces association of Smad2 and Smad3 (ref. 13). We examined the effect of Smad6 on the heteromerization of these SMAD proteins. Smad2 formed a complex with Smad4 in the presence of constitutively active T β R-I, as shown by coprecipitation of Smad4 with Smad2. The complex formation was abolished by Smad6 (Fig. 4a, top). However, the interaction of Smad3 and Smad4 induced by T β R-I was not affected by Smad6 (Fig. 4b), consistent with the finding that Smad6 does not inhibit Smad3 phosphorylation (Fig. 3c). Furthermore, heteromerization of Smad2 and Smad3 was inhibited by Smad6 (Fig. 4c), suggesting that phosphorylation of both proteins is necessary for this interaction. These results suggest that Smad6 interferes specifically with the activation of Smad2 in TGF- β signalling.

We tested the role of Smad6 in TGF- β signalling in luciferase reporter gene assays. p3TP-Lux, a sensitive reporter for TGF- β , was used in R mutant mink cells deficient in T β R-I (Fig. 5). Wild-type T β R-I restored the TGF- β response in these cells. Smad6 suppressed the activation of the reporter gene in a dose-dependent manner (Fig. 5a) and also suppressed transcriptional activation by constitutively active T β R-I (Fig. 5b). Expression of cyclin A is necessary for cell-cycle progression and is suppressed by TGF- β ¹⁴. Smad6 counteracted TGF- β in the cyclin A luciferase assay (Fig. 5c). These results indicate that Smad6 interferes with TGF- β signals in two distinct responses.

TGF- β is the prototype of a large family of cytokines that are involved in various biological activities¹⁻³. Molecules related to TGF- β thus act in environments in which multiple signals interact and are likely to be under tight spatial and temporal regulation. SMAD proteins are required for signalling in the TGF- β superfamily, and could be targets for other forms of control inside the cell. Smad1, Smad2, Smad3 and Smad5 transduce ligand-specific signals, whereas Smad4 is an essential common partner of these ligand-specific SMAD proteins. Our results show that Smad6 belongs to a third class of the SMAD family. It has quite a different structure from the other SMAD proteins, and is likely to be a negative regulator of signalling by the TGF- β superfamily. □

Methods

Cloning of mouse Smad6 and northern blot analysis. A mouse lung cDNA library (Stratagene) was screened with an EST clone (429356) as a probe. One of the clones contained the entire coding region of mouse Smad6 and was sequenced using an ALFred sequencer (Pharmacia Biotech) and a Sequenase sequencing kit (USB). Sequence analysis was done with DNASTAR. Human and mouse tissue blots (Clontech) were probed with the EST clone.

Plasmids. Mammalian expression vectors with an N-terminal tag (Flag or Myc) were constructed by inserting oligonucleotides encoding the epitope-tag sequence into pcDNA3 (Invitrogen). The coding region of the mouse Smad6 was amplified by the polymerase chain reaction (PCR) and subcloned into

Myc-pcDNA3 or Flag-pcDNA3. The integrity of the products was confirmed by sequencing. Smad1, Smad2, Smad3 and Smad4 expression plasmids were constructed in a similar manner.

Affinity crosslinking and immunoprecipitation. Iodination of TGF- β 1 (R&D Systems), activin A (gift from Y. Eto) and OP-1/BMP-7 (gift from T. K. Sampath) and the following immunoprecipitation were performed as described¹⁵.

Western blot and *in vivo* phosphorylation. COS-7 cells were transiently transfected using DMR1E-C (GibcoBRL). [³²P]orthophosphate- or [³⁵S]methionine/cysteine-labelling and immunoprecipitation were done as described¹⁶. For western blot analysis of the immunoprecipitated proteins, tagged proteins were detected by chemiluminescence (ECL, Amersham).

Luciferase assays. Mink R mutant cells were transiently transfected with an appropriate combination of a reporter, expression plasmids, and pcDNA3 using Tfx-50 (Promega). Total amounts of transfected DNA were the same in each experiment, and values were normalized using sea-pansy luciferase activity under the control of the thymidine kinase promoter (pRL-TK, Toyo, Ink).

Received 19 May; accepted 16 July 1997.

- Massagué, J., Hata, A. & Liu, F. TGF- β signalling through the Smad pathway. *Trends Cell Biol.* **7**, 187–192 (1997).
- Roberts, A. B. & Sporn, M. B. in *Peptide Growth Factors and Their Receptors* Part 1 (eds Sporn, M. B. & Roberts, A. B.) 419–472 (Springer, Heidelberg, 1990).
- Yingling, J. M., Wang, X.-F. & Basing, C. H. Signaling by the transforming growth factor- β receptors. *Biochim. Biophys. Acta* **1242**, 115–136 (1995).
- Liu, F. et al. A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**, 620–623 (1996).
- Hoodless, A. et al. MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* **85**, 489–500 (1996).
- Kretschmar, M. et al. The TGF- β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev.* **11**, 984–995 (1997).
- Suzuki, A. et al. Smad5 induces ventral fates in *Xenopus* embryo. *Dev. Biol.* **184**, 402–405 (1997).
- Eppert, K. et al. MADR2 maps 18q21 and encodes a TGF- β -regulated MAD-related protein that is functionally mutated in colorectal carcinoma. *Cell* **86**, 543–552 (1996).
- Zhang, Y., Feng, X.-H., Wu, R.-Y. & Derynck, R. Receptor-associated Mad homologues synergize as effectors of the TGF- β response. *Nature* **383**, 168–172 (1996).
- Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massagué, J. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**, 832–836 (1996).
- Macias-Silva, M. et al. MADR2 is a substrate of the TGF- β receptor and its phosphorylation is required for nuclear accumulation and signalling. *Cell* **87**, 1215–1224 (1996).
- Riggins, G. J. et al. Mad-related genes in the human. *Nature Genet.* **13**, 347–349 (1996).
- Nakao, A. et al. Heteromeric complex formation and activation of Smad2, Smad3 and Smad4 in TGF- β receptor-mediated signalling. *EMBO J.* (in the press).
- Feng, X.-H., Filvaroff, E. H. & Derynck, R. Transforming growth factor- β (TGF- β)-induced down-regulation of cyclin A expression requires a functional TGF- β receptor complex. *J. Biol. Chem.* **270**, 24237–24245 (1995).
- Okadome, T. et al. Characterization of the interaction of FKBP12 with the transforming growth factor- β type I receptor *in vivo*. *J. Biol. Chem.* **271**, 21687–21690 (1996).
- Nakao, A. et al. Identification of Smad2, a human Mad-related protein in the transforming growth factor β signaling pathway. *J. Biol. Chem.* **272**, 2896–2900 (1997).

Acknowledgements. We thank A. Nakao for the Smad1 and Smad2 cDNAs; S. Kern for the Smad4/DPC4 cDNA; J. Massagué for the p3TP-Lux reporter and R mutant cells; R. Derynck for Smad3 cDNA and the pCAL2 cyclin A reporter; Y. Eto for activin A; T. K. Sampath for OP-1/BMP-7; P. Horton for critical reading of the manuscript; and Y. Inada for technical assistance. This work was supported by grants-in-aid for scientific research from the Ministry of Education, Science, Sports and Culture of Japan, and special coordination funds for promoting science and technology from the Science and Technology Agency. K.M. is supported by the Toray Scientific Foundation and Uehara Memorial Foundation.

Correspondence and requests for materials should be addressed to M.K. (e-mail: mkawabat-ind@umin.u-tokyo.ac.jp). The DNA sequence of mouse Smad6 has been deposited in the GenBank database (accession no. AF010133).

Daughters against dpp modulates dpp organizing activity in *Drosophila* wing development

Kazuhide Tsuneizumi, Takuya Nakayama*, Yuko Kamoshida, Thomas B. Kornberg†, Jan L. Christian* & Tetsuya Tabata

Institute of Molecular and Cellular Biosciences, The University of Tokyo, Yayoi 1-1-1, Tokyo 113, Japan

* Department of Cell and Developmental Biology, Oregon Health Sciences University, Portland, Oregon 97201, USA

† Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA

The family of TGF- β signalling molecules play inductive roles in various developmental contexts¹. One member of this family, *Drosophila* Decapentaplegic (Dpp)² serves as a morphogen that patterns both the embryo^{3,1} and adult^{4,5}. We have now isolated a gene, *Daughters against dpp* (*Dad*), whose transcription is induced by Dpp. *Dad* shares weak homology with *Drosophila* Mad (Mothers against dpp)⁶, a protein required for transduction of Dpp signals. In contrast to Mad or the activated Dpp receptor, whose overexpression hyperactivates the Dpp signalling pathway, overexpression of *Dad* blocks Dpp activity. Expression of *Dad* together with either Mad or the activated receptor rescues phenotypic defects induced by each protein alone. *Dad* can also antagonize the activity of a vertebrate homologue of Dpp, bone morphogenetic protein (BMP-4; ref. 7), as evidenced by induction of dorsal or neural fate following overexpression in *Xenopus* embryos. We conclude that the pattern-organizing mechanism governed by Dpp involves a negative-feedback circuit in which Dpp induces expression of its own antagonist, *Dad*. This feedback loop appears to be conserved in vertebrate development.

The *Drosophila* wing is divided into two compartments along its anteroposterior (A/P) axis. The compartment boundary between these regions serves as the source of an organizing activity that patterns both anterior and posterior compartments^{8–10}. This activity is mediated, at least in part, by the long-range action of Dpp^{8,11}, which is expressed by cells along the A/P compartment boundary (Fig. 1a). Dpp is thought to act as a morphogen to inform target cells of their position along the A/P axis^{4,5}, but as little is known about how cells interpret the distribution of Dpp protein, we conducted an enhancer trap screen to identify genes whose transcription is controlled by Dpp. We identified two enhancer trap lines in the same locus (89E/F), P1883 and 1(3)1E4, whose expression patterns are similar to those of Dpp during embryonic (data not shown) and imaginal development (Fig. 1b). We designate the gene whose expression is reflected in these enhancer traps *Daughters against dpp* (*Dad*). In these enhancer trap lines, β -galactosidase is expressed in a wide stripe that straddles the A/P compartment boundary of the imaginal discs, in contrast to Dpp, whose expression is confined to the anterior side. This pattern of expression suggests that *Dad* expression is positively regulated by the secreted Dpp molecule. To test whether *Dad* responds to Dpp signalling, we examined its expression in P1883 wing discs in which a *UAS-dpp* transgene is transcribed in a ring around a wing pouch under the control of a Gal 4 driver¹² (Fig. 1f). Ectopic Dpp expression resulted in abnormally large discs and in ectopic expression of *Dad* in a broad ring around a wing pouch (Fig. 1d). Identical results were obtained when another transgene, *UAS-tkv*^{Q253D} (ref. 5), which encodes a constitutively active form of the major type-I Dpp receptor, Thick

veins (Tkv), was used (Fig. 1e). In addition, expression of *Dad* was not detected in cells that lacked a functional Tkv Dpp receptor (Fig. 1g–i). These results indicate that Dpp signalling is necessary and sufficient for *Dad* expression in the developing wing.

To isolate the *Dad* gene, genomic DNA surrounding the insertion of 1(3)1E4 was cloned by plasmid rescue. An exon-containing genomic fragment was identified by *in situ* (Fig. 1c) and northern hybridization and was used as a probe to isolate complementary DNAs. The longest cDNA is 3.8 kb and contains a putative open reading frame capable of encoding a 63.6K, 568-residue protein. The deduced amino-acid sequence shows limited homology to *Drosophila* Mad⁶, a protein that is required for intracellular transduction of the Dpp signal (Fig. 2). A family of structurally related proteins, termed SMAD proteins, has been identified in *Caenorhabditis elegans*¹³ and in vertebrates¹⁴. It has been postulated that these proteins are phosphorylated in response to receptor activation, after which they translocate to the nucleus and function as transcription factors. SMAD proteins share conserved amino- and carboxy-terminal domains separated by a variable proline-rich region. Although *Dad* shares significant homology with other SMAD family members within the carboxy-terminal domain, the

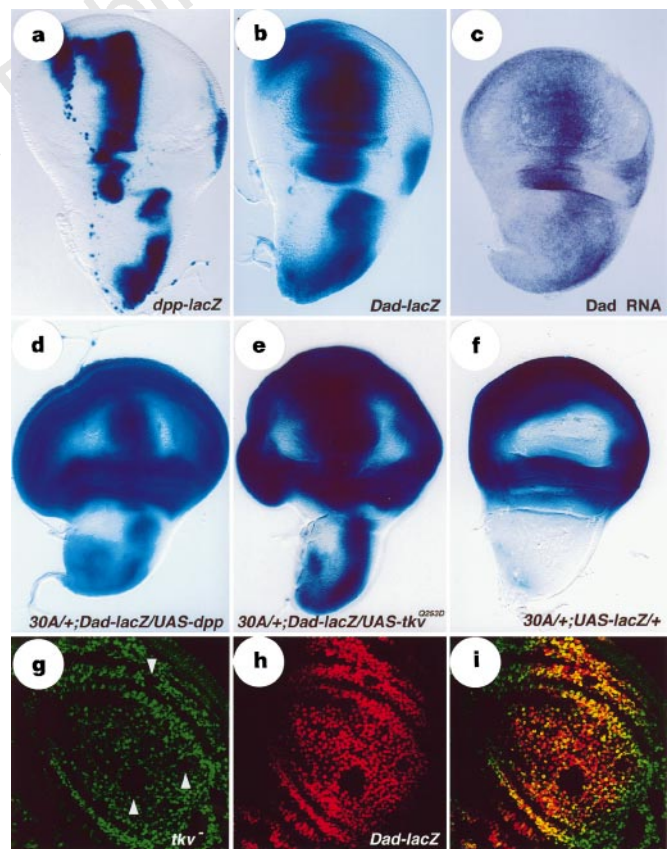


Figure 1 *Dad* expression is positively controlled by Dpp in the wing imaginal disc. **a–c**, Wild-type expression pattern of Dpp revealed by the enhancer trap P1552³¹ (**a**) and *Dad* revealed by the enhancer trap P1883 (**b**) or *in situ* hybridization (**c**) with a digoxigenine-labelled genomic fragment. The stripe of *Dad*-expressing cells is broader than that of Dpp. **d–f**, *Dad* expression, as monitored by staining for β -galactosidase activity in P1883, changes correspondingly when Dpp (**d**) or the activated Dpp receptor, Tkv^{Q253D} (**e**) is ectopically expressed under the control of the Gal4 driver 30A. **f**, *UAS-lacZ* expression in strain 30A. **g–i**, *Dad* expression requires Dpp signalling. Clones of cells mutant for *tkv*, identified by the absence of Myc antigen (**g**, green) did not express *Dad* (**h**) which was monitored by immunostaining against β -galactosidase (red) in P1883. **i**, Superimposed images of **g** and **h**. Mutant clones are marked by arrowheads. All discs shown in this and subsequent figures are wing discs of third instar larvae. Posterior is to the right and ventral is up.

amino-terminal domain is less well conserved. The carboxy-terminal domain of Dad shares the highest homology with human Smad6 (ref. 15), which lacks an amino-terminal homology domain (Fig. 2).

To analyse its function, we ectopically expressed Dad using the Gal4–UAS system and assayed for patterning defects in the wing. When Dad was misexpressed along the wing margin, the wing lost its margin and, in extreme cases, only a tiny winglet formed (Fig. 3b). We obtained similar results with four independent lines and with several different drivers, although the degree of the phenotypic defects varied from line to line (data not shown). Because *dpp* is

required not only for pattern formation, but also for cell proliferation, clones of cells that have lost *Dpp* responsiveness as a result of mutation of genes encoding *tkv* or *punt* (a type II *Dpp* receptor) do not survive in the wing blade¹⁶. The similarity in wing phenotype caused by loss of *Dpp* responsiveness and that caused by ectopic expression of Dad suggest that Dad antagonizes the *Dpp* signalling pathway. In contrast, conditions that mimicked the effects of hyperactivating the *Dpp* signalling pathway, such as overexpressing *Tkv*^{Q253D} or *Mad*, cause outgrowth of wing tissue (Fig. 3c, d). When Dad was expressed together with either *Tkv*^{Q253D} or *Mad*, phenotypic defects caused by overexpression of each protein alone were

Figure 2 Deduced amino-acid sequence of Dad and alignment with *Drosophila* Mad⁹, human Smad1 (ref. 32) and Smad6 (ref. 15). Residues identical in more than two sequences are indicated with a black background and residues conserved only between Dad and Smad6 are shaded.

MAD	1	-----M
SMAD1	1	-----M
SMAD6	1	-----M
DAD	1	MIFPREKKVLRVYASNNPNSNGVSAAPPAQPPRPPPPHRRPRPHQCTPSPGYSCNRSDSL
MAD	2	DTDDVESNTS S AMSTLG S IPSPSTSPAVKRL L GWKQG --- DEEEKWAERAVD S LVRKLL
SMAD1	1	-----M
SMAD6	1	-----M
DAD	61	AMRQTPLPPY S SIACGMD C SS S SSCGQS S LS Q GGHNNNSHPYRRLPNHMD L LP
MAD	57	RR K GA E EL E AL S CPG Q PS K CV T IP S LD --- GR L Q V SR K GL P H V IT C RV R ---
SMAD1	4	-----M
SMAD6	1	-----M
DAD	121	PSACDRCC T AP G V E CA S SCD D ML D GS D L D QDRSPD Q Q V P V D R MI S ATT T TM F R K C
MAD	108	-----W P D L Q S HE L K P LE L C Q Y P FS A R K Q E VC I NP Y HY K R --- V
SMAD1	94	-----M
SMAD6	1	-----M
DAD	181	CGGATESTSGSTLT I P V ST E RATA H P P Q T Q A Q G K R PRE D PE L M K Q L R R Q R NE L L A V
MAD	145	ESPVLPPVLVPRHSE F AP G HS M L Q F N H V A --- EP S MP H NP V S S NG ---
SMAD1	131	-----M
SMAD6	1	-----M
DAD	241	K E RL D P E FT K T Q R D VVR P TT T TA P TY L Q C IL I P C K T Q T V M EP H VT A SL R L P WR R L N AK L
MAD	188	PNSESLSTNTSVG --- SP S -----VN S MP H NP V S S NG ---
SMAD1	179	-----M
SMAD6	12	-----M
DAD	301	K R L P TC A AN D C I Y M C N PL H W F R L H Q E T SE P TP P Y Q R S K N L R K D AP E ED S Q N D A K
MAD	231	-----C S -----NN P ND G Q L L D A Q MG D V A Q V ST S EP F W A SI A Y E L L C R VE G EP F C
SMAD1	235	-----M
SMAD6	47	-----M
DAD	361	SAALS W SAR S TS S IS E Y K PA L YES V TT D G --- K D HN I NS Q V N Q L HA M EB M AR V GE F PA E
MAD	280	NN S V L VD Q FT R -----PS I NS D RC C COL S RV --- NR S T I ET R RR I CG V ET T Y V CE S VA
SMAD1	290	-----M
SMAD6	89	-----M
DAD	419	XT N -----AV N Y T Y D GI V ASE V D S MC I RD L TP A GN Q IS V VT A RT V GE V LS E RG D PN I
MAD	336	E C LS D SA I FP V Q S R N C N Y H RG F PS T V C K I FP G CS L K I FP N Q E FA Q LL S Q S V N NG F E A V E
SMAD1	346	-----M
SMAD6	142	-----M
DAD	477	RR R NT T IP V D S -----SEN --- LD R V C RY M PG Y CD R AP T -----NR A EL L MR D H G H P NG
MAD	396	LT R K T IR S MF V RG W GA E Y R OD V TS T PC W IE I EL G PL W LD K VL T OM G SP H AI S SV S
SMAD1	406	-----M
SMAD6	197	-----M
DAD	530	RY D FN S V I TS F AK G WG F C V S R OP T TS Q PC W IE I EL G PL W LD K VL T OM G SP H AI S SV S

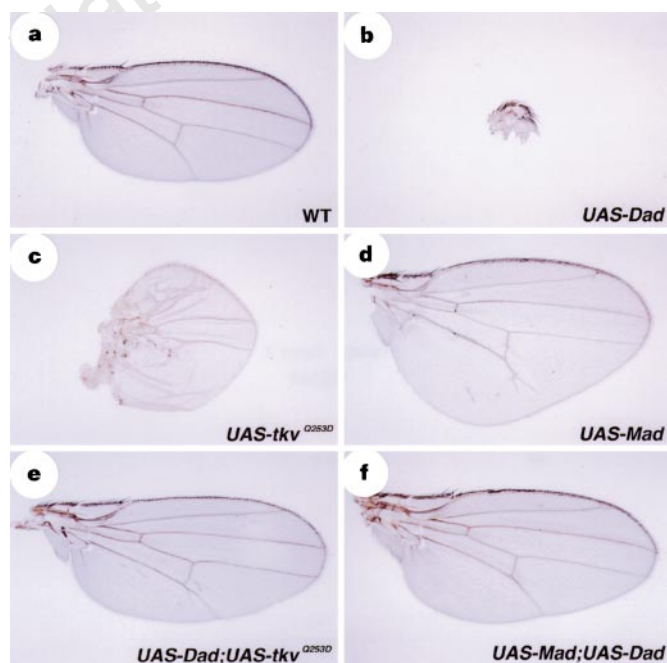


Figure 3 Functional interactions between *Dad*, *tkv*, and *Mad* in adult wings. **a**, Wild type. **b**, Small winglet produced when Dad was expressed ectopically under the control of *vg-Gal4* along the prospective wing margin. **c**, General overproliferation after expression of *Tkv*^{Q253D} under the control of *vg-Gal4*. **d**, Overproliferation of the P compartment after ectopic expression of *Mad* under the control of *vg-Gal4*. **e**, Adult wings appeared normal when both Dad and *Tkv*^{Q253D} were overexpressed under the control of *vg-Gal4*. **f**, Phenotypic rescue after ectopic expression of both Dad and *Mad* under the control of *vg-Gal4*. Phenotype was variable, from nearly wild type (**f**) to resembling shapes characteristic of the phenotype of overexpression of *Mad* alone, as in (**d**).

nullified and nearly wild-type wings formed (Fig. 3e, f).

We then examined the effects of ectopic expression of Dad and Mad on a Dpp target gene, *optomotor-blind* (*omb*)¹⁷ which is positively regulated by *dpp* in wing discs. *Omb* is expressed in a broad stripe straddling Dpp expressing cells, as is Dad, but unlike Dad, it is not expressed along the entire A/P boundary. Clones of cells that expressed Dad or Mad, or both, had effects that were cell autonomous. *Omb* expression was absent in Dad-overexpressing cells (Fig. 4a–c). When Mad-overexpressing clones fell inside, or near the normal *Omb*-expression domain, higher level, or ectopic expression of *Omb* was observed, respectively. In contrast, when Mad-overexpressing cells were situated distal to the endogenous *Omb*-expression domain, ectopic expression of *Omb* was not

detected (Fig. 4d–f). Thus, overexpressed Mad may be dependent on Dpp signalling for activation whereas elevated levels of Mad may lower the threshold concentration of Dpp required to induce *Omb* expression. When both Dad and Mad were overexpressed in the same cells, *Omb* expression was barely affected (Fig. 4g–i). The observation that overexpression of Dad did not reduce expression of endogenous Dpp (data not shown), together with the fact that patterning defects induced by overexpression of Dad were rescued by overexpression of Mad or Tkv^{Q253D}, suggest that antagonism of the Dpp signal by Dad occurs after reception of the Dpp signal at the cell surface and before control of transcription of target genes.

To confirm that Dad represses *Omb* expression, we analysed somatic *Dad* mutant clones. *Dad* mutants were induced by excising the P element of the 1(3)1E4 enhancer trap line; one of these, *Dad*²⁷¹⁻⁶⁸, is associated with a deletion of the entire C-terminal domain after amino acid 391. *Omb* was derepressed autonomously in *Dad*²⁷¹⁻⁶⁸ clones in wing discs (Fig. 4j–l), indicating that Dad normally represses *Omb* expression. We infer that Dad negatively modulates the level of Dpp signalling, at least during wing development.

Components of the Dpp signalling pathway are highly conserved between arthropods and chordates⁷, so we investigated whether Dad can antagonize signalling by BMP-4, a vertebrate homologue of Dpp. In embryos of *Xenopus laevis*, BMP-4 functions to specify ventral mesodermal and epidermal fates. Blockage of BMP signalling during gastrulation induces the formation of secondary dorsal axes in intact embryos and neuralizes the fate of explanted ectodermal cells⁷. As shown in Fig. 5a–d, similar patterning defects were observed following ectopic expression of Dad in *Xenopus* embryos. Microinjection of Dad mRNA into dorsal blastomeres of 4-cell embryos produced no detectable patterning defects (Fig. 5b), whereas injection into ventral cells induced formation of a secondary axis in 90% of embryos (Fig. 5c). The induced axes contained muscle (Fig. 5c, middle panel) and neural tissue, but not notochord (Fig. 5c, right). In some cases, a cyclopic eye differentiated in the secondary axis (Fig. 5c, right), although the frequency of eye induction varied, ranging from 3–38% in different experiments. Ectodermal explants (animal caps) from Dad-injected embryos elongated and formed darkly pigmented cement glands (Fig. 5d, arrowheads in lower left panel), whereas explants from control embryos retained a rounded epidermal appearance. Dad-injected explants expressed cement gland- (*XAG*) and neural-specific (*N-CAM*, *OtxA*) genes but not a dorsal mesodermal gene, α -actin or a panmesodermal gene, *Xbra* (Fig. 5d, right), indicating that Dad can directly mediate neural induction in the absence of mesoderm. To test whether Dad can antagonize BMP function, we co-injected BMP-4 and Dad mRNAs into dorsal blastomeres of 4 cell embryos (Fig. 5e). Injection of BMP-4 mRNA alone caused a loss of anterior and dorsal structures, yielding an average dorsoanterior index (DAI; ref. 18) of 1.2 ($n = 30$), whereas co-injection of Dad mRNA almost completely rescued the ventralized phenotype (average DAI = 4.7, $n = 39$). These results suggest that Dad can antagonize BMP signalling in *Xenopus* embryos. Given that components of the Dpp/BMP signalling pathway are highly conserved between insects and vertebrates, it is likely that a homologue of Dad exists that modulates the amplitude or duration of BMP signalling during vertebrate embryogenesis.

Although Dad is a distantly related member of the SMAD family, it is unique among SMADs in antagonizing, rather than transducing, TGF- β -like signals. Oddly enough, Dad appears to participate in a direct negative feedback loop in that it antagonizes the very signalling pathway (that is, Dpp) that is required for induction of its own expression. This relationship between Dpp, which positively regulates the level of expression of Dad, and Dad, which negatively regulates the level of Dpp activity, suggests that the final outcome of Dpp signalling may not be directly proportional to the graded concentration of Dpp protein but may depend on the balance between transduction of Dpp signals by activated Mad, and antag-

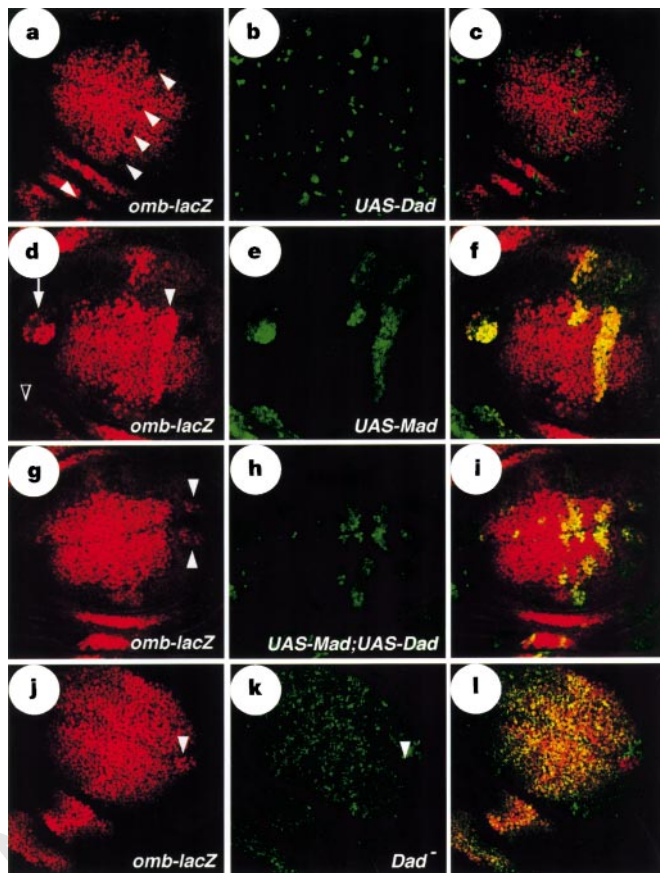


Figure 4 Dad and Mad effects on *Omb* expression. Clones of cells that overexpressed Dad or Mad or both under the control of *actin-Gal4* were generated by the flip-out technique. *Omb* expression was revealed by antibody staining against β -galactosidase in wing discs of the *omb*^{P1} enhancer strain. *Omb* expression (a, red) was lost in clones of cells that ectopically expressed high levels of Dad (revealed by the presence of GFP, green, b). c, Superimposed images of a and b. d, Mad-expressing clones within (arrowhead) or near (arrow) the normal *Omb*-expressing domain express higher than normal levels of *Omb* (red). However, Mad-expressing cells distal to the normal *Omb*-expressing domain, did not express *Omb* ectopically (d, open arrowhead). Mad-expressing cells were marked by the presence of GFP (e, green). f, Superimposed images of d and e. g, Clones of cells that overexpressed both Dad and Mad did not detectably alter *omb* expression. Some clones near the normal *omb*-expressing domain (arrowhead) variably expressed low levels of *Omb*. Cells expressing both Dad and Mad were marked by the presence of GFP (h, green). i, Superimposed images of g and h. j–l, Clones of cells mutant for *Dad* ectopically express *Omb*. *Omb* (j, red) was ectopically expressed in clones of cells mutant for *Dad* (arrowhead), identified by the absence of Myc antigen (k, green). l, Superimposed images of j and k. Mutant clones are marked by an arrowhead. Clones were generated 48 to 72 h (d–l) or 72 to 96 h (a–c) after egg-laying and *Omb* expression was observed 24 h (a–c) or 48 h (d–l) later.

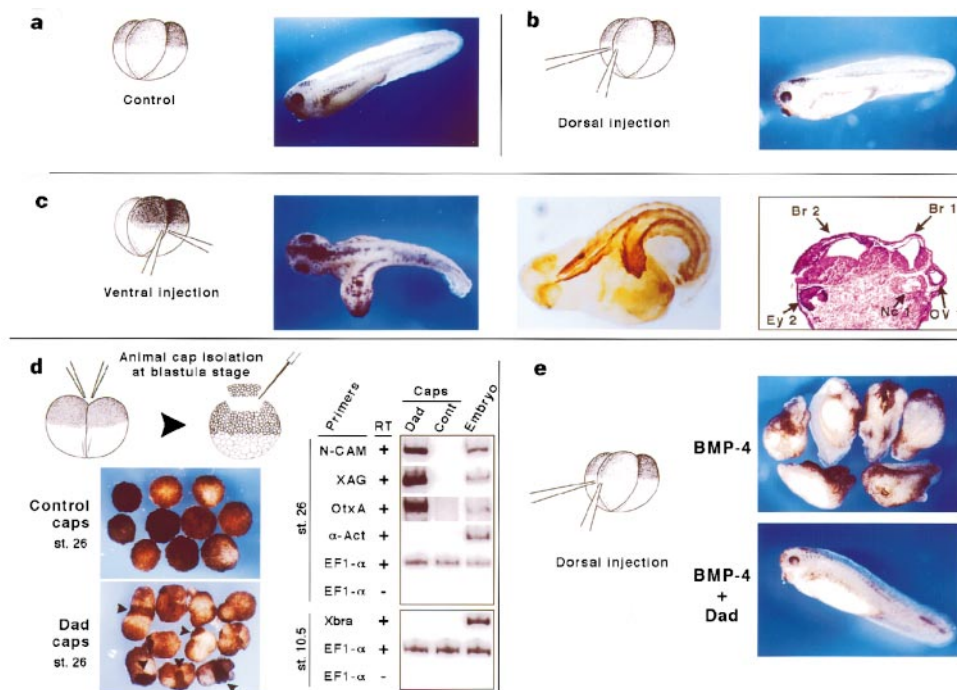


Figure 5 Dad antagonizes BMP-4 activity in *Xenopus* embryos. **a–c**, Dad induces a partial secondary axis in *Xenopus* embryos. 200 pg synthetic RNA encoding Dad was injected into two blastomeres near the dorsal or ventral marginal zone of 4 cell embryos as shown schematically. Embryos injected dorsally (**b**) ($n = 69$; 7% showed nonspecific abnormality) are indistinguishable from uninjected sibling embryos (**a**) ($n = 216$; 4% showed nonspecific abnormality) whereas 90% of embryos injected ventrally developed with a partial secondary axis (**c**) ($n = 86$ from 4 independent experiments; 10% showed nonspecific abnormality) in which muscle was detected by immunostaining with 12/101 antibody (middle panel). A transverse section (**c**, right panel) through the anterior end of a twinned embryo shows the presence of brain and eye (Br2 and Ey2, respectively) and absence of notochord in the secondary axis, and the presence of brain, otic vesicle and notochord (Br1, OV1, Nc1, respectively) in the primary axis. **d**, Dad directly

induces neural tissue in the absence of mesoderm in ectodermal explants (animal caps). Dad mRNA (200 pg) was injected near the animal pole of 2-cell embryos as shown, and animal caps were explanted at the blastula stage and cultured in isolation until sibling embryos reached stage 10.5 or stage 26. Dad-injected caps (Dad caps, lower left panel) form cement glands (arrowheads), whereas control caps (upper left panel) remain epidermal. Explants were analysed by reverse transcription with the polymerase chain reaction (RT-PCR) for expression of neural- (OtxA, N-CAM), cement gland- (XAG) or mesodermal- (α -actin, Xbra)-specific gene expression. Dad-injected caps expressed neural, but not mesodermal marker genes. **e**, Dad antagonizes BMP-mediated ventralization. Dorsal injection of BMP-4 mRNA (500 pg) leads to a loss of all dorsal and anterior structures; co-injection of Dad mRNA (250 pg) completely rescues the ventralized phenotype.

onism of Dpp signals by Dad. Mechanistically, Dad may interact directly with Mad to modulate Dpp signalling because co-expression of Dad and Mad rescue phenotypic defects induced by either protein alone. Precedence for direct interaction between different members of the SMAD family exists in that the SMADs form multimeric complexes¹⁹. Alternatively, Dad may antagonize signalling by interacting directly with the receptors. A vertebrate SMAD protein, Smad2, stably associates with receptors and blocks TGF β -dependent transcriptional responses when its conserved three C-terminal serine residues are substituted with alanine residues²⁰. In contrast, wild-type Smad2 transiently associates with receptors and transduces the signal. Dad does not have these C-terminal serines and may stably associate with receptors. A new member of SMADs, Smad7, which does not have these C-terminal serines, was reported to inhibit TGF β signalling by interacting stably with the receptor²¹. Our data suggest the existence of a Dad negative feedback circuit that might stabilize the gradient of positional information emanating from Dpp expressing cells. □

Methods

Cloning of the Dad gene. Genomic DNA adjacent to the P-element transposon in strain 1(3)1E4 was isolated by plasmid rescue. A 3.7-kb fragment of the rescued plasmid was used to isolate genomic clones from a λ phage library (gift from Y.-N. Jan). Genomic fragments containing exons were identified by northern blotting and fragments that hybridized to RNA species with a distribution identical to the expression pattern of β -galactosidase in 1(3)1E4 and P1883 were used to screen a 4–8-h embryonic cDNA library from N. Brown. Three independent cDNA clones were sequenced on both strands

using an automated sequencer.

Ectopic expression. Transgenes and the enhancer trap line used were *UAS-dpp* and *vestigial-Gal4* (gift from S. Morimura), *UAS- tkv^{Q253D}* (ref. 5), *UAS-Mad*²², and *omb*^{P1} (ref. 17). *UAS-Dad* was constructed with a 3.5 kb fragment containing the entire coding region inserted into pUAST¹². Discs from the larvae of the genotype *30A/+; P1883/UAS-dpp* (Fig. 1d), *30A/+; P1883/UAS- tkv^{Q253D}* (Fig. 1e) and *30A/+; UAS-lacZ/+* (Fig. 1f) were stained for β -galactosidase activity. Wings from the adults of the genotype *vg-Gal4/UAS-Dad* (Fig. 3b), *vg-Gal4/+; UAS- $tkv^{Q253D}/+$* (Fig. 3c), *UAS-Mad/+; vg-Gal4/+* (Fig. 3d), *vg-Gal4/UAS-Dad; UAS- $tkv^{Q253D}/+$* (Fig. 3e) and *UAS-Mad/+; vg-Gal4/UAS-Dad* (Fig. 3f) were examined. Dad- or Mad-expressing clones (Fig. 4) were made by combination of a FLP-out technique and Gal4/UAS system²³. Briefly, excision of FLP-out cassette by FLP from *AyGal4* places the Gal4 coding sequence just downstream to the actin 5C promoter which had been separated by the FLP-out cassette from the Gal4 coding sequence. Discs from larvae of the genotype *y w h s FLP omb^{P1}/y w; AyGal4 UAS-GFP/UAS-Dad* (Fig. 4a–c), *y w h s FLP omb^{P1}/y w UAS-Mad; AyGal4 UAS-GFP/+* (Fig. 4d–f), and *y w h s FLP omb^{P1}/y w UAS-Mad; AyGal4 UAS-GFP/UAS-Dad* (Fig. 4g–i) were immunostained for lacZ expression. The Gal4-expressing clones also expressed UAS-GFP. **Clonal analysis.** Clones of cells mutant for *tkv* were made with the FLP-FRT technique. Discs from *y w h s FLP/+; tkvⁿ¹² FRT40/ π Myo FRT40; P1883/+* larvae were immunostained for β -galactosidase and Myc protein. *tkvⁿ¹² FRT40* was a gift from K. Basler. Clones of cells mutant for Dad were made by γ -ray irradiation (1,350 rads from a ¹³⁷Cs source). Discs from *y w omb^{P1}/+; FRT82 π Myo/Dad²⁷¹⁻⁶⁸* larvae were immunostained for β -galactosidase and Myc proteins.

Xenopus embryo culture and manipulation. *Xenopus* eggs were obtained and embryos microinjected and cultured as described²⁴. Histological analysis,

Identification of Smad7, a TGF β -inducible antagonist of TGF- β signalling

Atsuhito Nakao*, Mozghan Afrakhte†, Anita Morén*, Takuya Nakayama‡, Jan L. Christian‡, Rainer Heuchel*, Susumu Itoh*, Masahiro Kawabata§, Nils-Erik Heldin†, Carl-Henrik Heldin* & Peter ten Dijke*

*Ludwig Institute for Cancer Research, Box 595, Biomedical Center, S-751 24 Uppsala, Sweden

†Department of Pathology, University Hospital, S-751 85 Uppsala, Sweden

‡Department of Cell and Developmental Biology, L215, Oregon Health Sciences University, School of Medicine, 3181 SW Sam Jackson Park Road, Portland, Oregon 97201-3098, USA

§Department of Biochemistry, The Cancer Institute, Tokyo, The Japanese Foundation for Cancer Research, and Research for the Future Program, Japan Society for the Promotion of Science, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan

TGF- β signals from the membrane to the nucleus through serine/threonine kinase receptors and their downstream effectors, termed SMAD proteins¹. The activated TGF- β receptor induces phosphorylation of two such proteins, Smad2 and Smad3 (refs 2–6), which form hetero-oligomeric complex(es) with Smad4/DPC4 (refs 5–10) that translocate to the nucleus^{2,4,5,7}, where they then regulate transcriptional responses^{11,12}. However, the mechanisms by which the intracellular signals of TGF- β are switched off are unclear. Here we report the identification of Smad7, which is related to Smad6 (ref. 13). Transfection of Smad7 blocks responses mediated by TGF- β in mammalian cells, and injection of Smad7 RNA into *Xenopus* embryos blocks activin/TGF- β signalling. Smad7 associates stably with the TGF- β receptor complex, but is not phosphorylated upon TGF- β stimulation. TGF β -mediated phosphorylation of Smad2 and Smad3 is inhibited by Smad7, indicating that the antagonistic effect of Smad7 is exerted at this important regulatory step. TGF- β rapidly induces expression of Smad7 mRNA, suggesting that Smad7 may participate in a negative feedback loop to control TGF- β responses.

Smad7 was identified as an expressed sequence tag (EST) related to known SMAD proteins. Sequence analysis of isolated mouse and human cDNAs predict that mouse Smad7 and human Smad7 have 426 amino-acid residues with 98% identity (Fig. 1a). Smad7 is most related to Smad6 (ref. 13), with 36% and 56% sequence identities in the amino-terminal domain and the carboxy-terminal Mad homology (MH)2 domain, respectively. The Smad7 amino-terminal domain shows only weak similarity to the MH1 domains found in Smad1 to Smad5. RNA blot analysis with a Smad7 probe revealed one main transcript of approximately 4.4 kilobases (Fig. 1b). Among the tissues analysed, the highest expression of Smad7 was found in the lung.

In order to investigate whether Smad7 modulates the responsiveness to TGF- β , we transfected the TGF β -inducible p3TPLux luciferase reporter construct, which contains the PAI-1 promoter, into Mv1Lu mink epithelial cells in the absence or presence of Smad7 cDNA. Smad7 was found to inhibit TGF β 1-induced luciferase activity (Fig. 2a). Moreover, the induction of p3TPLux response by constitutively active variants of the TGF- β receptor T β R-I and a structurally related type IB receptor for activin (ActR-IB), when transfected in R-mutant cells, was also inhibited by co-transfection with Smad7 (Fig. 2b, and data not shown). Thus our results indicate that Smad7 is a potent negative regulator of p3TPLux response induced by both T β R-I and ActR-IB. In addition, Smad7, but not

immunostaining of whole embryos (12/101 antibody was obtained from the Developmental Studies Hybridoma Bank, NICHD), and RT-PCR analysis of RNA extracted from cultured animal caps was done as described²⁵ except that 14- μ m sections were cut and PCR cycles were as follows: 95 °C for 5 min followed by a variable number of cycles (determined empirically for each primer pair) at 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s. Control reactions in which reverse transcriptase was omitted were run in parallel for all samples. The sequences of *EF1- α* , *N-CAM*²⁶, α -actin²⁷, *OtxA*²⁸, *XAG1* (ref. 29); *Xbra*³⁰ primers have been reported previously.

Received 27 May; accepted 20 August 1997.

- Hogan, B. L. M. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev.* **10**, 1580–1594 (1996).
- Padgett, R. W., St Johnston, R. D. & Gelbart, W. M. A transcript from a *Drosophila* pattern gene predicts a protein homologous to the transforming growth factor- β family. *Nature* **325**, 81–84 (1987).
- Ferguson, L. F. & Anderson, K. V. Decapentaplegic acts as a morphogen to organize dorsal–ventral pattern in the *Drosophila* embryo. *Cell* **71**, 451–461 (1992).
- Lecuit, T. et al. Two distinct mechanisms for long-range patterning by Decapentaplegic in the *Drosophila* wing. *Nature* **381**, 387–393 (1996).
- Nellen, D., Burke, R., Struhl, G. & Basler, K. Direct and long-range action of a DPP morphogen gradient. *Cell* **85**, 357–368 (1996).
- Sekelsky, J., Newfeld, S., Raftery, L., Chartoff, E. & Gelbart, W. Genetic characterization and cloning of *Mothers against dpp*, a gene required for decapentaplegic function in *Drosophila melanogaster*. *Genetics* **139**, 1347–1358 (1995).
- Ferguson, E. L. Conservation of dorsal–ventral patterning in arthropods and chordates. *Curr. Opin. Genet. Dev.* **6**, 424–431 (1996).
- Zecca, M., Basler, K. & Struhl, G. Sequential organizing activities of *engrailed*, *hedgehog* and *decapentaplegic* in the *Drosophila* wing. *Development* **121**, 2265–2278 (1995).
- Tabata, T., Schwartz, C., Gustavson, E., Ali, Z. & Kornberg, T. B. Creating a *Drosophila* wing *de novo*, the role of *engrailed*, and the compartment border hypothesis. *Development* **121**, 3359–3369 (1995).
- Lawrence, P. A. & Struhl, G. Morphogens, compartments, and pattern: lessons from *Drosophila*? *Cell* **85**, 951–961 (1996).
- Capdevila, J. & Guerrero, I. Targeted expression of the signaling molecule *decapentaplegic* induces pattern duplications and growth alterations in *Drosophila* wings. *EMBO J.* **13**, 4459–4468 (1994).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Savage, C. et al. *Caenorhabditis elegans* genes *smad-2*, *smad-3* and *smad-4* define a conserved family of transforming growth factor- β pathway components. *Proc. Natl Acad. Sci. USA* **93**, 790–794 (1996).
- Massague, J., Hata, A. & Fang, L. TGF- β signalling through the Smad pathway. *Trends Cell Biol.* **7**, 187–192 (1997).
- Riggins, G. J. et al. Mad-related genes in the human. *Nature Genet.* **13**, 347–349 (1996).
- Burke, R. & Basler, K. DPP receptors are autonomously required for cell proliferation in the entire developing *Drosophila* wing. *Development* **122**, 2261–2269 (1996).
- Grimm, S. & Pflugfelder, G. O. Control of the gene optomotor-blind in *Drosophila* wing development by *decapentaplegic* and *wingless*. *Science* **271**, 1601–1604 (1996).
- Kao, K. R. & Elinson, R. P. Dorsalization of mesoderm induction by lithium. *Dev. Biol.* **132**, 81–90 (1989).
- Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massague, J. Partnership between DPC4 and SMAD in TGF- β -signalling pathways. *Nature* **383**, 832–836 (1996).
- Macias-Silva, M. et al. MADR2 is a substrate of the TGF β receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell* **87**, 1215–1224 (1997).
- Hayashi, H. et al. The Mad-related protein Smad7 associates with the TGF β receptor and functions as an antagonist of TGF β signaling. *Cell* **89**, 1165–1173 (1997).
- Newfeld, S. J., Chartoff, E. H., Graff, J. M., Melton, D. A. & Gelbart, W. M. *Mothers against dpp* encodes a conserved cytoplasmic protein required in DPP/TGF- β responsive cells. *Development* **122**, 2099–2108 (1996).
- Ito, K., Awano, W., Suzuki, K., Hiromi, Y. & Yamamoto, D. The *Drosophila* mushroom body is a quadruple structure of clonal units each of which contains a virtually identical set of neurones and glial cells. *Development* **124**, 761–771 (1997).
- Moon, R. T. & Christian, J. L. Microinjection and expression of synthetic mRNAs in *Xenopus* embryos. *Technique* **1**, 76–89 (1989).
- Cui, Y., Tian, Q. & Christian, J. L. Synergistic effects of Vg1 and Wnt signals in the specification of dorsal mesoderm and endoderm. *Dev. Biol.* **180**, 22–34 (1996).
- Kengaku, M. & Okamoto, H. bFGF as a possible morphogen for the anteroposterior axis of the central nervous system in *Xenopus*. *Development* **121**, 3121–3130 (1995).
- Hemmati-Brivanlou, A. & Melton, D. A. Inhibition of activin receptor signaling promotes neuralization in *Xenopus*. *Cell* **77**, 273–281 (1994).
- Lai, C.-J., Ekker, S. C., Beachy, P. A. & Moon, R. T. Patterning of the neural ectoderm of *Xenopus laevis* by the amino-terminal product of hedgehog autoproteolytic cleavage. *Development* **121**, 2349–2360 (1995).
- Blitz, I. L. & Cho, K. W. Y. Anterior neuroectoderm is progressively induced during gastrulation: the role of the *Xenopus* homeobox gene *orthodenticle*. *Development* **121**, 993–1004 (1995).
- Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* **77**, 283–295 (1994).
- Tabata, T. & Kornberg, T. B. Hedgehog is a signalling protein with a key role in patterning in *Drosophila*. *Cell* **76**, 89–102 (1994).
- Liu, F. et al. A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**, 620–623 (1996).

Acknowledgements. We thank I. Yabe for help with the confocal microscopy, H. Eguchi for help with the γ -ray irradiation, and K. Basler, Y. Hiromi, Y.-N. Jan, S. Morimura, S. Newfeld, G. Pflugfelder, A. Spradling and the Bloomington *Drosophila* Stock Centre for fly strains. This work was supported by grants from the Japan Society for the Promotion of Science (Research for the Future Program), from the New Energy and Industrial Technology Development Organization, and the Uehara Memorial Foundation to T.T., by grants-in-aid from the Ministry of Education, Science, Sports and Culture of Japan to T.T. and K.T., by grants from the NIH to J.L.C. and T.B.K., and from the American Heart Association to J.L.C. T.N. was the recipient of a postdoctoral fellowship from the Toyobo Biotechnology Foundation (Japan) and is currently supported by the American Heart Association, Oregon Affiliate.

Correspondence and requests for materials should be addressed to T.T. (e-mail: ttataba@hgc.ims.u-tokyo.ac.jp). DDBJ sequence accession number is AB004232.

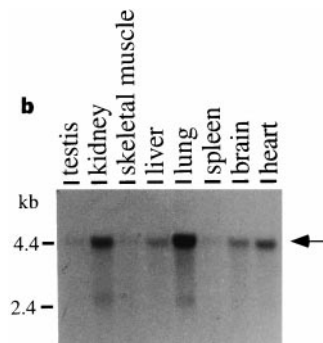
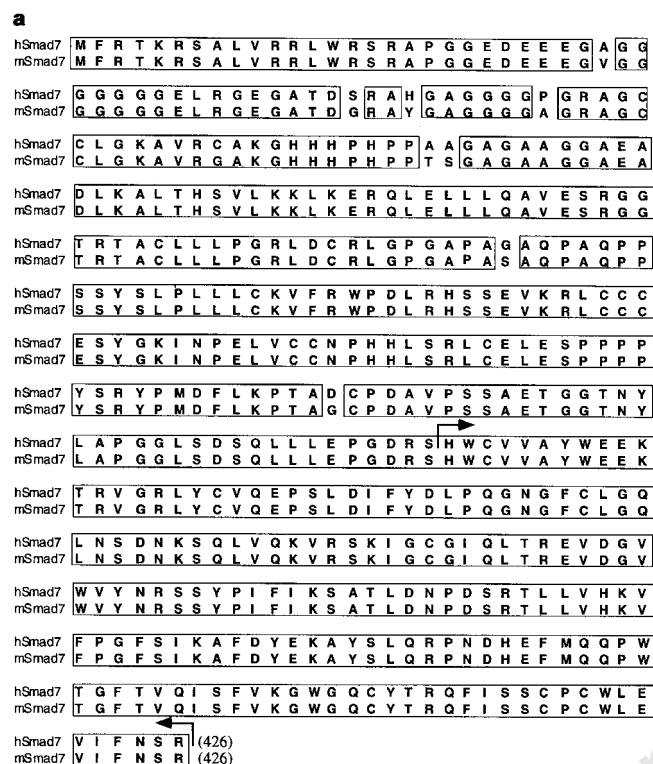


Figure 1 Amino-acid sequence of Smad7 and expression of Smad7 mRNA. **a**, Predicted amino-acid sequence of mouse (m) and human (h) Smad7. The borders of the Mad-homology (MH)2 domain are indicated by arrows. Genbank accession numbers for mouse and human Smad7 are AF015260 and AF015261, respectively. **b**, Expression of Smad7 mRNA in various mouse tissues. A blot with mRNAs from mouse tissues was hybridized with an N-terminal domain Smad7 probe.

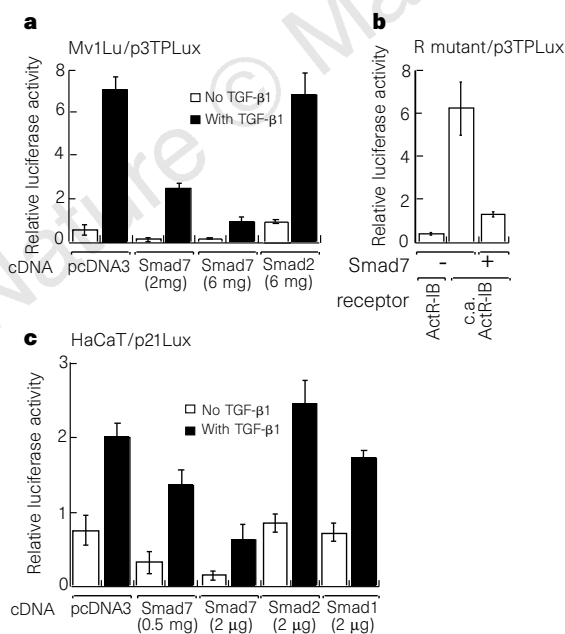


Figure 2 Effects of mouse Smad7 transfection on responses induced by TGF- β family members, and effects of RNA injection in embryos. **a**, Transfection of Smad7, but not Smad2, in Mv1Lu cells blocks the TGF β -induced p3TPLux response. **b**, Transfection of Smad7 in R mutant cells inhibits p3TPLux response induced by constitutively active (c.a.) ActR-IB. **c**, Transfection of Smad7, but not Smad 2 or Smad1, in HaCaT cells antagonizes the TGF β -induced p21 Lux response. **d-g**, Injection of 500 pg synthetic RNA encoding either Smad7 or myc-epitope tag (MT)²⁹ into each blastomere of two-cell embryos, as shown schematically (**d**); embryos were cultured to the tailbud stage. Embryos injected with MT

RNA developed normally (top embryo in each panel), whereas embryos made to misexpress Smad7 (bottom embryo in each panel) failed to form head or tail structures (**e**) and showed a complete or partial loss of muscle (**f**) and notochord (**g**). Arrows indicate immunoreactive muscle (**f**) and notochord (**g**) in MT-injected embryos. **h, i**, Injection of 500 pg synthetic RNA encoding MT (**h**) or Smad7 (**i**) into one blastomere of two-cell embryos near the equator. Expression of *brachyury* was analysed by whole-mount *in situ* hybridization. MT-injected embryos show the typical ring of *brachyury* expression (**h**), but *brachyury* transcripts are not detected on one side of Smad7-injected embryos (**i**, arrow).

Smad1 or Smad2, was found to antagonize the TGF β -induced transcriptional response with a luciferase reporter construct containing the p21 CDK inhibitor promoter (p21Lux)¹³ in human keratinocytes (HaCaT) (Fig. 2c). Smad7 and N-terminally Flag-tagged Smad7 (F-Smad7) gave essentially the same antagonistic results on TGF β signalling, suggesting that N-terminal tagging does not interfere with Smad7 function. Thus Smad7 inhibits TGF β -induced pathways leading to extracellular matrix production as well as growth inhibition.

To determine whether Smad7 can inhibit TGF β /activin-like signals *in vivo*, we analysed patterning defects caused by over-expression of Smad7 in *Xenopus* embryos. When the endogenous activin signalling pathway is inactivated in early *Xenopus* embryos, by introduction of dominant-negative forms of either an activin receptor^{14,15} or of Smad4 (ref. 6), mesoderm fails to form. Similarly, microinjection of RNA encoding Smad7 into both blastomeres of two-cell embryos inhibited mesoderm formation. Specifically, head and tail structures were absent or severely deficient in 80% of injected embryos ($n = 168$; Fig. 2e), as were mesodermal derivatives such as muscle ($n = 60$; Fig. 2f) and notochord ($n = 48$; Fig. 2g). To test further the ability of Smad7 to block mesoderm formation *in vivo*, we analysed embryos for expression of *brachyury*, a gene that is expressed throughout the presumptive mesoderm during gastrulation¹⁶ (Fig. 2h). Injection of Smad7 RNA into the equatorial region of one blastomere of two-cell embryos prevented *brachyury* expression on one side of the embryo (Fig. 2i).

Two members of the TGF β family, activin and Vg1, are candidates for being endogenous mesoderm-inducing molecules¹⁷. We used a *Xenopus* animal-cap assay¹⁷ to test directly the ability of Smad7 to block signalling downstream of these ligands. Both activin and Vg1 induced expression of *brachyury* in ectodermal explants (animal caps), but coexpression of Smad7 inhibited *brachyury*

induction by 60% (data not shown), demonstrating that Smad7 can block activin- and Vg1-mediated mesoderm induction.

To investigate the mechanism by which Smad7 exerts its negative role in signalling by TGF β family members, we tested whether Smad7, like Smad2 and Smad3 (refs 3–5), can associate with the TGF β receptor complex. COS cells transfected with F-Smad7 in combinations with T β R-II (wild-type or kinase-deficient mutant) and T β R-I (wild-type or kinase-deficient mutant) were affinity-labelled with ¹²⁵I-TGF β 1, and cell lysates were immunoprecipitated with Flag antiserum. We found that Smad7 interacted very efficiently with the TGF β receptor complex (Fig. 3). Smad7 interacted with wild-type T β R-I as well as kinase-inactive T β R-I in complex with wild-type T β R-II. In contrast, Smad2 and Smad3 interact stably only with complex of kinase-deficient T β R-I and wild-type T β R-II (refs 4, 5). No interaction was observed between Smad7 and a heteromeric complex of kinase-deficient T β R-II and T β R-I, or with T β R-II alone (Fig. 3). Thus transphosphorylation of T β R-I by T β R-II kinase is required for the association of Smad7 with T β R-I.

We investigated whether Smad7 is phosphorylated upon association with TGF β receptors. COS cells were transfected with Smad7, or with Smad2 for comparison, in the absence or presence of receptors, labelled with [³²P]orthophosphate, and treated with TGF β 1. As expected, Smad2 phosphorylation was increased upon coexpression with receptors, and addition of TGF β 1 led to a further increase^{2,4–6}. However, we observed no (or only very weak) phosphorylation of Smad7 in cells transfected with constitutively active forms of T β R-I and ActR-IB, or in cells transfected with wild type T β R-I and T β R-II after stimulation with ligand (Fig. 4a). Thus, despite its direct association with T β R-I, Smad7 is not a substrate for the TGF β R-I kinase. Notably, Smad7 lacks the conserved SS(M/V)S motif in its carboxy tail (Fig. 1a), which in the case of Smad2 (ref. 4) and Smad1 (ref. 10) are phosphorylated by type-I receptor kinase.

Smad7 might exert its negative role in TGF β signalling by interfering with activation of other SMAD proteins. We therefore examined whether Smad7 affects the phosphorylation of Smad2 and Smad3 by using [³²P]orthophosphate-labelled COS cells transfected with TGF β receptors and SMAD proteins. Smad7 inhibited the TGF β 1-induced phosphorylation of Smad2 (Fig. 4b, c) and Smad3 (Fig. 4d, e). A Smad7 level twice as high as that of Smad3 was sufficient to cause a significant decrease in TGF β -induced receptor-mediated phosphorylation of Smad3. As previously reported^{4,5}, we found that Smad2 and Smad3 become phosphorylated when over-expressed in COS cells, which for Smad2 does not occur at the SSMS motif in the C-terminal tail⁴. Smad7 appears not to block this receptor-independent phosphorylation. Because activation of Smad2 and Smad3 is essential for optimal TGF β signalling^{3–6,10}, the inhibition of their phosphorylation provides a mechanistic explanation for the antagonistic effect of Smad7. The association of Smad7 with T β R-I suggests that it may compete with Smad2 and Smad3 for receptor binding. In accordance with this notion, co-transfection of Smad2 or Smad3 with Smad7 reduced the antagonistic effect of Smad7 in a TGF β 1-induced p3TPLux assay (data not shown). It has recently been reported that Smad7 inhibits TGF β signalling responses¹⁸, which is in agreement with our findings. We also found that Smad7 not only inhibits TGF β and activin signalling, but also blocks signalling by bone morphogenetic proteins (BMPs). Smad7 associated with BMPR-Is and inhibited phosphorylation of Smad1 in cultured mammalian cells (A.N., unpublished results). In addition, injection of Smad7 RNA into ventral cells of *Xenopus* embryos phenocopies the effect of blocking the BMP signalling pathway, and led to formation of an incomplete secondary dorsal axis (J.L.C., unpublished results).

Regulation of Smad7 expression by signalling molecules may provide a possible means of effectively modulating TGF β responses. We therefore investigated whether the expression of Smad7, and for comparison Smad2, Smad3 and Smad4, were

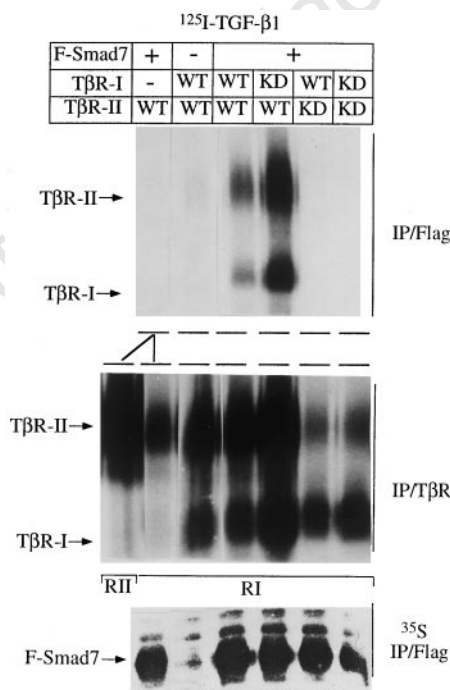


Figure 3 Association of Smad7 with the TGF β receptor complex. COS cells were transfected with F-Smad7 in combinations of wild-type (WT) or kinase-deficient (KD) mutants of T β R-II and T β R-I. The receptors were covalently affinity labelled with ¹²⁵I-TGF β 1. Immunoprecipitates were analysed by SDS-PAGE and FujiX Bio-Imager. Expression of receptors and F-Smad7 after transfection was determined by immunoprecipitation (IP) with specific antisera on lysates from cells affinity-labelled with ¹²⁵I-TGF β 1 or [³⁵S]methionine/cysteine-labelled transfected cells, respectively.

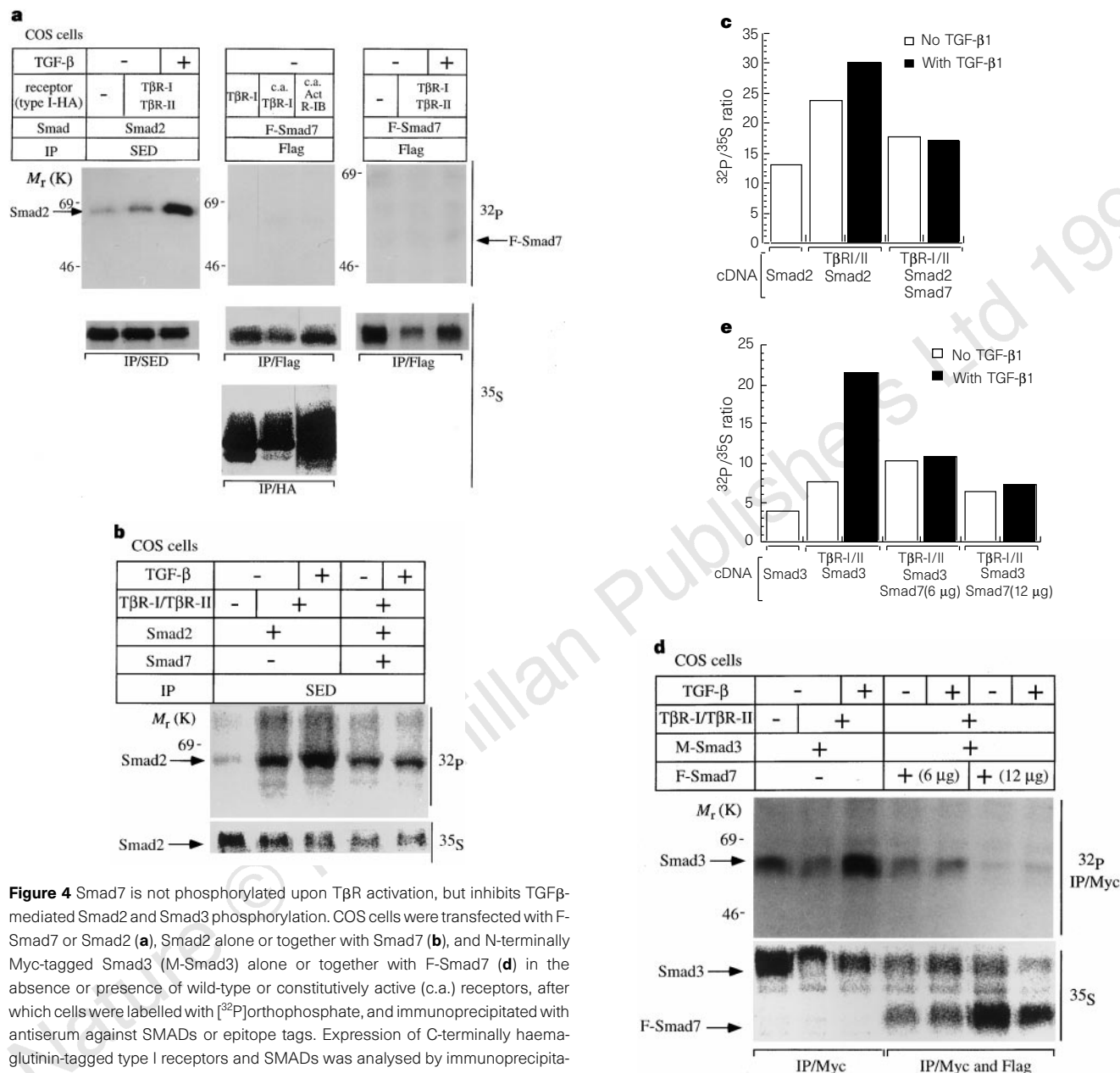


Figure 4 Smad7 is not phosphorylated upon T β R activation, but inhibits TGF β -mediated Smad2 and Smad3 phosphorylation. COS cells were transfected with F-Smad7 or Smad2 (**a**), Smad2 alone or together with Smad7 (**b**), and N-terminally Myc-tagged Smad3 (M-Smad3) alone or together with F-Smad7 (**d**) in the absence or presence of wild-type or constitutively active (c.a.) receptors, after which cells were labelled with [32 P]orthophosphate, and immunoprecipitated with antiserum against SMADs or epitope tags. Expression of C-terminally haemagglutinin-tagged type I receptors and SMADs was analysed by immunoprecipitation on portions of cell lysates, which had been labelled with [35 S]methionine/cysteine. The 32 P or 35 S radioactivity associated with Smad2 in **b** or Smad3 in **d** was quantified by using a Fujix Bio-Imager, and the 32 P/ 35 S ratio was calculated (**c** and **e**, respectively). Representative results of three independent experiments are shown.

regulated by TGF- β 1. Northern blot analysis of SMADs on RNA from TGF β 1-stimulated Mv1Lu cells, SW1736 human anaplastic thyroid carcinoma cells, and HaCat cells, revealed that Smad7 mRNA, but not Smad2, Smad3 or Smad4 mRNA, was rapidly induced in response to TGF- β 1 stimulation (Fig. 5a, b). In Mv1Lu cells, Smad7 mRNA was induced fourfold more after 30 min of TGF- β 1 stimulation. Induction of Smad7 mRNA by TGF- β 1 was observed in the presence of cycloheximide (CHX), indicating that *de novo* protein synthesis is not required for this response. Smad7 mRNA was actually superinduced in the presence of TGF- β 1 and CHX (Fig. 5b), probably as a result of an increase in mRNA stability or loss of transcriptional repressors by CHX. Treatment with CHX alone only slightly increased the basal Smad7 mRNA level (data not shown). In contrast, the phorbol ester TPA (12-*O*-tetradecanoylphorbol-13-acetate), which like TGF- β 1 stimulates PAI-1 expression and inhibits the growth of

Mv1Lu cells, did not induce Smad7 expression (Fig. 5c). Taken together, these data indicate that Smad7 is an immediate-early response gene for TGF- β 1, and may act in a negative feedback loop to regulate the intensity or duration of the TGF- β signal.

TGF- β -like signalling responses have been shown to be regulated at both the ligand level^{19,20} and the receptor level^{21–23}. Our results indicate that another important level of modulation of TGF- β signals may be exerted through intracellular agonistic and antagonistic SMADs. Loss of expression or overexpression of Smad7 may contribute to diseases in which TGF- β has been implicated. □

Methods

Isolation of mSmad7 and hSmad7 cDNA and northern blot analysis. cDNA encoding the complete mSmad7 was made by fusing mouse EST cDNA (AA061644) with a partial cDNA isolated from a mouse placenta library. The cDNA for hSmad7 was isolated by screening a human brain cDNA library.

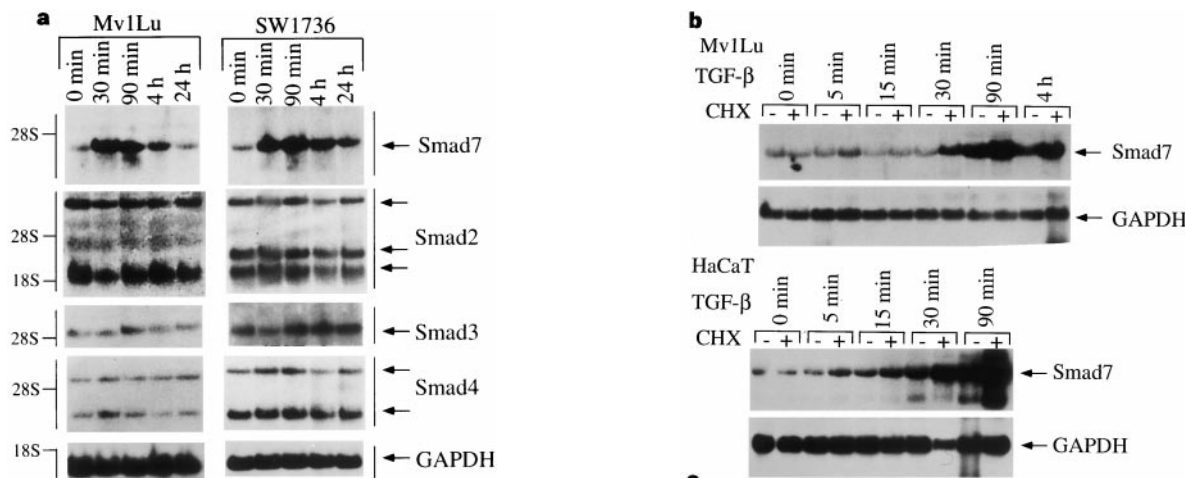


Figure 5 Northern blot analyses. **a**, Northern blot analysis of Smads on RNA from Mv1Lu cells and SW1736 cells on exposure to TGF- β 1 (10 ng ml⁻¹). **b**, Northern blot analysis of Smad7 mRNA using total RNA isolated from Mv1Lu cells and HaCaT cells in the absence or presence of CHX (20 μ g ml⁻¹) on TGF- β 1 stimulation. CHX was added 30 min before TGF- β 1. **c**, Northern blot analysis of RNA from Mv1Lu cells showing that PAI-1 mRNA, but not Smad7, is induced on TPA (10 nM) stimulation. The amount of total RNA loaded (20 μ g per lane) was checked by hybridization with a GAPDH probe.

Isolation of total RNA and northern blot analysis were performed essentially as described²⁴.

Expression plasmids and antisera. Expression constructs for T β R-II, T β R-I, T β R-II, ActR-IB, Smad1, Smad2 and Smad3, and antisera recognizing Smads or receptors, have been described⁵. F-Smad7 was made by PCR and subcloning into pcDNA3-Flag.

Cell assays. Transfection of COS cells, metabolic labelling, immunoprecipitation, affinity crosslinking, [³²P]orthophosphate labelling of cells, and SDS-PAGE were performed as described⁵.

Transcriptional response assay. Mv1Lu and R mutant cells were transfected with p3TPLux as described⁵. HaCaT cells were transiently transfected with p21 Lux²⁵ using Transfectam reagent from Promega (Madison, WI). In each experiment, total equal amounts of DNA were transfected. Luciferase activity was measured as described⁵, and normalized for transfection efficiency. Results shown are representative of at least four independent experiments.

Xenopus embryo culture and manipulation. *Xenopus* eggs were obtained and embryos microinjected and cultured as described²⁶. For animal-cap assays, 200 pg of RNA encoding activin- β B or a BMP-Vg1 chimera²⁷ was injected either alone or together with 400 pg of Smad7 RNA. Immunostaining of whole embryos (12/101 antibody was obtained from the Developmental Studies Hybridoma Bank under contract N01-HD-2-3144 from the NICHD; Tor70 was obtained from R. Harland), whole-mount *in situ* hybridization, RT-PCR analysis of RNA extracted from cultured animal caps, and quantification of *brachyury* expression relative to that of *EF-1a* were performed as described^{16,28}.

Received 14 May; accepted 8 September 1997.

- Massagué, J., Hata, A. & Liu, F. TGF- β signalling through the Smad pathway. *Trends Cell Biol.* **7**, 187–192 (1997).
- Eppert, K. *et al.* MADR2 maps to 18q21 and encodes a TGF β -regulated MAD-related protein that is functionally mutated in colorectal carcinoma. *Cell* **86**, 543–552 (1996).
- Zhang, Y., Feng, X.-H., Wu, R.-Y. & Derynck, R. Receptor-associated Mad homologues synergize as effectors of the TGF- β response. *Nature* **383**, 168–172 (1996).
- Macías-Silva, M. *et al.* MADR2 is a substrate of the TGF β receptor and its phosphorylation is required for nuclear accumulation and signalling. *Cell* **87**, 1215–1224 (1996).
- Nakao, A. *et al.* TGF- β receptor mediated signalling through Smad2, Smad3 and Smad4. *EMBO J.* **16**, 5353–5362 (1997).
- Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massagué, J. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**, 832–836 (1996).
- Zhang, Y., Musci, T. & Derynck, R. The tumor suppressor Smad4/DPC4 as a central mediator of Smad function. *Curr. Biol.* **7**, 270–276 (1997).
- Wu, R.-Y., Zhang, Y., Feng, X.-H. & Derynck, R. Heteromeric and homomeric interactions correlate with signalling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol. Cell. Biol.* **17**, 2521–2528 (1997).

- Shi, Y., Hata, A., Lo, R. S., Massagué, J. & Pavletich, N. P. A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature* **388**, 87–93 (1997).
- Kretschmar, M., Liu, F., Hata, A., Doody, J. & Massagué, J. The TGF- β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev.* **11**, 984–995 (1997).
- Chen, X., Rubock, M. J. & Whitman, M. A transcriptional partner for MAD proteins in TGF- β signalling. *Nature* **383**, 691–696 (1996).
- Kim, J., Johnson, K., Chen, H. J., Carroll, S. & Laughon, A. *Drosophila* Mad binds to DNA and directly mediates activation of *vestigial* by Decapentaplegic. *Nature* **388**, 304–308 (1997).
- Imamura, T. *et al.* Smad6 inhibits signalling by the TGF- β superfamily. *Nature* **389**, 622–626 (1997).
- Hemmati-Brivanlou, A. & Melton, D. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in *Xenopus* embryos. *Nature* **359**, 609–614 (1992).
- Chang, C., Wilson, P. A., Mathews, L. S. & Hemmati-Brivanlou, A. A. *Xenopus* type I activin receptor mediates mesodermal but not neural specification during embryogenesis. *Development* **124**, 827–837 (1997).
- Smith, J. C., Price, B. M. J., Green, J. B., Weigel, D. & Herrman, B. G. Expression of a *Xenopus* homolog of *Brachyury* (*T*) is an immediate-early response to mesoderm induction. *Cell* **67**, 79–87 (1991).
- Kessler, D. S. & Melton, D. A. Vertebrate embryonic induction: mesodermal and neural patterning. *Science* **266**, 596–604 (1994).
- Hayashi, H. *et al.* The MAD-related protein Smad7 associates with the TGF β receptor and functions as an antagonist of TGF β signaling. *Cell* **89**, 1165–1173 (1997).
- Zimmerman, L. B., De Jesus-Escobar, J. M. & Harland, R. M. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein-4. *Cell* **86**, 599–606 (1996).
- Piccolo, S., Sasai, Y., Lu, B. & De Robertis, E. M. Dorsal patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* **86**, 589–598 (1996).
- Wang, T. *et al.* The immunophilin FKBP12 functions as a common inhibitor of the TGF β family type I receptors. *Cell* **86**, 435–444 (1996).
- Chen, Y.-G., Liu, F. & Massagué, J. Mechanism of TGF β receptor inhibition by FKBP12. *EMBO J.* **16**, 3866–3876 (1997).
- Luo, K. & Lodish, H. F. Positive and negative regulation of type II TGF- β receptor signal transduction by autophosphorylation on multiple serines. *EMBO J.* **16**, 1970–1981 (1997).
- Afrakhte, M., Nister, M., Ostman, A., Westermarck, B. & Paulsson, Y. Production of cell-associated PDGF-AA by a human sarcoma cell line: evidence for a latent autocrine effect. *Int. J. Cancer* **68**, 802–809 (1996).
- Datto, M. B., Yu, Y. & Wang, X. F. Functional analysis of the transforming growth factor β responsive elements in the WAF/Cip/p21 promoter. *J. Biol. Chem.* **270**, 28623–28628 (1995).
- Moon, R. T. & Christian, J. L. Microinjection and expression of synthetic mRNAs in *Xenopus* embryos. *Technique* **1**, 76–89 (1989).
- Dale, L., Matthews, G. & Colman, A. Secretion and mesoderm-inducing activity of the TGF- β related domain of *Xenopus* Vg1. *EMBO J.* **12**, 4471–4480 (1993).
- Cui, Y., Tian, Q. & Christian, J. L. Synergistic effects of Vg1 and Wnt signals in the specification of dorsal mesoderm and endoderm. *Dev. Biol.* **180**, 22–34 (1996).
- Turner, D. & Weintraub, H. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8**, 1434–1447 (1994).

Acknowledgements. We thank K. Miyazono, T. Imamura and K. Cho for discussion; I. Nakao for encouragement; and H. Ohashi, R. Derynck, X.-F. Wang and J. Massagué for agents. This work was supported in part by grants from the NIH (J.L.C.), the American Heart Association, Oregon Affiliate (T.N.) and the Swedish Medical Research Council (N.-E.H.).

Correspondence and requests for materials should be addressed to P.T.D. (e-mail: Peter.ten_Dijke@LICR.uu.se).

Survival of FimH-expressing enterobacteria in macrophages relies on glycolipid traffic

David M. Baorto*, Zhimin Gao*, Ravi Malaviya*, Michael L. Dustin*, Anton van der Merwe†, Douglas M. Lublin* & Soman N. Abraham*‡§

Departments of * Pathology and ‡ Molecular Microbiology, Washington University School of Medicine, St Louis, Missouri 63110, USA
† MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK

Strains of *Escherichia coli* persist within the human gut as normal commensals, but are frequent pathogens and can cause recurrent infection^{1–3}. Here we show that, in contrast to *E. coli* subjected to opsonic interactions stimulated by the host's immune response, *E. coli* that bind to the macrophage surface exclusively through the bacterial lectin FimH can survive inside the cell following phagocytosis. This viability is largely due to the attenuation of intracellular free-radical release and of phagosome acidification during FimH-mediated internalization, both of which are triggered by antibody-mediated internalization. This different processing of non-opsonized bacteria is supported by morphological evidence of tight-fitting phagosomes compared with looser, antibody-mediated phagosomes. We propose that non-opsonized FimH-expressing *E. coli* co-opt internalization of lipid-rich microdomains following binding to the FimH receptor, the glycosylphosphatidylinositol-linked protein CD48, because (1) the sterol-binding agents filipin, nystatin and methyl β -cyclodextrin specifically block FimH-mediated internalization; (2) CD48 and the protein caveolin both accumulate on macrophage membranes surrounding bacteria; and (3) antibodies against CD48 inhibit FimH-mediated internalization. Our findings bring the traditionally extracellular *E. coli* into the realm of opportunistic intracellular parasitism and suggest how opportunistic infections with FimH-expressing enterobacteria could occur in a setting deprived of opsonizing antibodies.

Escherichia coli are normal residents of the alimentary canal but they have also been implicated in opportunistic and often recurrent infections in man and animals^{1–3}. *E. coli*, as well as many other enteric bacteria, express on their surface type-1 fimbriae, 1–2- μ m-long filaments composed of several subunits including FimH, a mannose-binding lectin that is responsible for promoting bacterial adherence and colonization of mucosal surfaces^{4,5}. Paradoxically, however, FimH also promotes adhesion to phagocytic cells, an event that should presumably result in early bacterial clearance. With two exceptions^{6,7}, most *in vitro* studies have reported that type-1 fimbrial-mediated association with phagocytic cells results in bacterial killing^{8–10}. As the effect of opsonic serum components or the contribution of other bacterial surface antigens that promote phagocytosis was not excluded, these findings may not be definite. Nevertheless, expression of FimH is viewed by some as being detrimental to bacterial virulence *in vivo*^{10,11}, an assumption that could not account for the invariable expression of type 1 fimbriae by *E. coli* isolated from clinical specimens^{12,13}.

To determine whether FimH-mediated binding to phagocytic cells is a fatal error for the bacterium, we developed a system directly to compare the fate of intracellular *E. coli* in mouse bone-marrow-

derived macrophages following internalization by a FimH-mediated, non-opsonic mechanism, to the fate of the same organism internalized by an antibody-mediated opsonic mechanism. After phagocytosis, extracellular bacteria were pulsed with enough gentamicin to eradicate them, allowing exclusive observation of intracellular *E. coli*. In contrast to antibody-internalized *E. coli*, which are effectively killed in the first hour, *E. coli* internalized by a FimH-mediated mechanism showed almost no sign of killing after

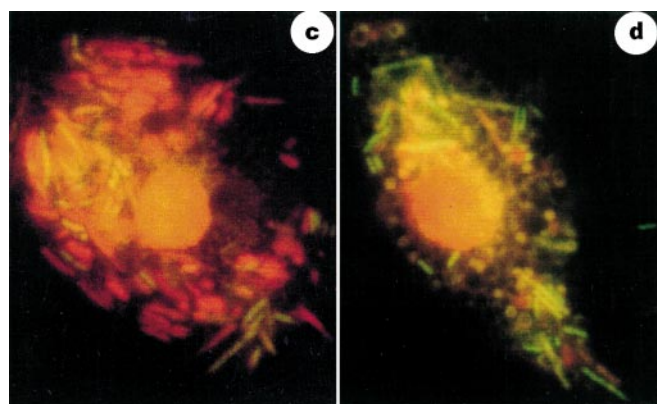
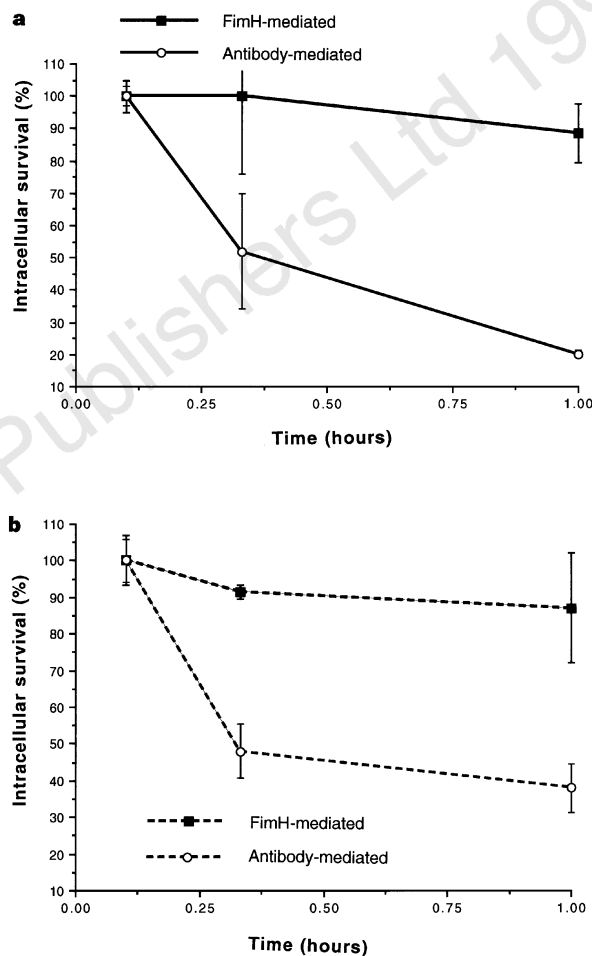


Figure 1 Intracellular survival of FimH-internalized *E. coli* in macrophages compared to opsonized *E. coli*. Intracellular *E. coli* can survive in cultured macrophages when they are internalized via FimH, in contrast to opsonized *E. coli* internalized via antibody, which are effectively killed (**a**, model strain; **b**, uropathogenic-clinical isolate). Bacterial viability stains also demonstrate increased intracellular survival of the FimH-internalized bacteria at 1 hour (**c**, antibody-internalized; **d**, FimH-internalized). Viable bacteria stain green and dead bacteria stain red.

§ Present address: Department of Pathology, Box 3712, Duke University Medical Center, Durham, North Carolina 27710, USA.

internalization (Fig. 1). This result was true for both the model strain (Fig. 1a) and a uropathogenic clinical isolate (Fig. 1b). The FimH-internalized *E. coli* were intracellularly persistent even after 24 hours, when all antibody-internalized *E. coli* were dead. Assessment of intracellular bacterial viability by vital staining in more heavily infected cultures indicates the diminished capacity of phagocytic cells to kill *E. coli* internalized non-opsonically (Fig. 1c, d). Thus, the fate of *E. coli* engulfed by fimbrial lectins is distinct from that mediated by bacteria-specific antibodies. When bacteria lacking FimH were coated with the glucose/mannose-binding plant lectin concanavalin A, intracellular survival rates were comparable

to that of opsonized FimH⁺ bacteria (data not shown), showing that the response of phagocytic cells to FimH was specific to this lectin.

Although the oxidative burst evoked in FimH-internalized *E. coli* is nearly equivalent to that evoked by opsonically internalized bacteria (Fig. 2a), the intracellular release of free radicals is remarkably attenuated in the case of FimH-mediated phagocytosis (Fig. 2b). The small difference between the FimH- and antibody-induced oxidative burst, as measured by luminol-enhanced chemiluminescence (Fig. 2a), may reflect the difference seen in the intracellular burst. *Bordetella pertussis* bacteria avoid death by promoting phagocytosis through the complement CR3 receptor, which abrogates release of toxic oxygen products¹⁴. Our findings indicate that macrophages have an intrinsic capacity selectively to discharge toxic oxygen metabolites intra- or extracellularly which can be modulated by *E. coli* FimH. In addition to the different oxidative burst pattern, we find that the intracellular compartment containing FimH-internalized bacteria is acidified less efficiently than that containing antibody-opsonized bacteria (Fig. 2c), which is reminiscent of the deficient acidification in *Mycobacterium* species and *Histoplasma capsulatum*^{15,16}, although the mode of entry was not implicated as in our case.

Transmission electron microscopy revealed that FimH-internalized *E. coli* are directed to an intracellular compartment that is also morphologically distinct from opsonized bacteria. About 85% of FimH-internalized bacteria were in a tight-fitting compartment after an hour, whereas more than 90% of antibody-internalized *E. coli* were in more spacious phagosomes with surrounding matrix material (Fig. 3a, b).

We investigated the molecular components on the macrophage responsible for diverting FimH-internalized bacteria through an apparently unique endocytic pathway. We have identified the putative membrane receptor on rat and mouse mast cells for FimH-expressing bacteria to be CD48, a glycoprotein moiety attached to the plasma membrane by a glycoposphatidylinositol (GPI)-linked anchor (R.M. *et al.*, manuscript in preparation).

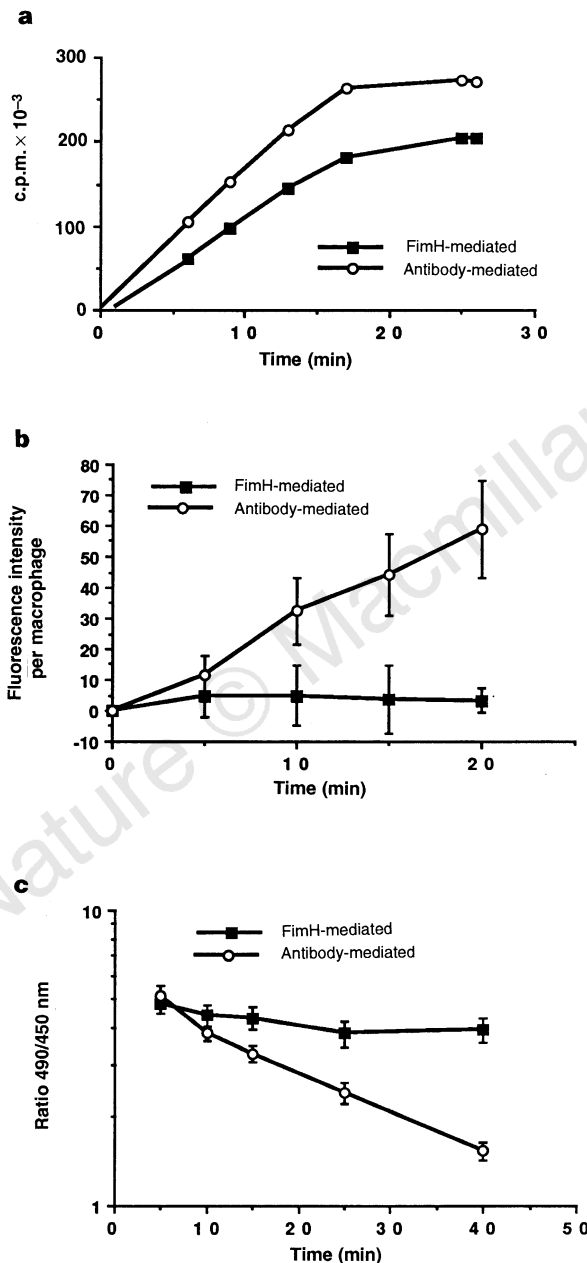


Figure 2 Attenuation of both oxidative and non-oxidative bactericidal mechanisms following FimH-mediated phagocytosis. Both FimH and antibody-mediated contact with macrophages induces an oxidative burst as detected by chemiluminescence (a). However, the intracellular component of the oxidative burst clearly induced during antibody-mediated phagocytosis is absent during FimH-mediated phagocytosis (b). The FimH-internalized *E. coli* are present in a compartment following phagocytosis that is acidified to a much lesser degree than the antibody-internalized *E. coli* (c). The ratio of FITC emission at 490/450 nm is proportional to the pH of the bacterial compartment²².

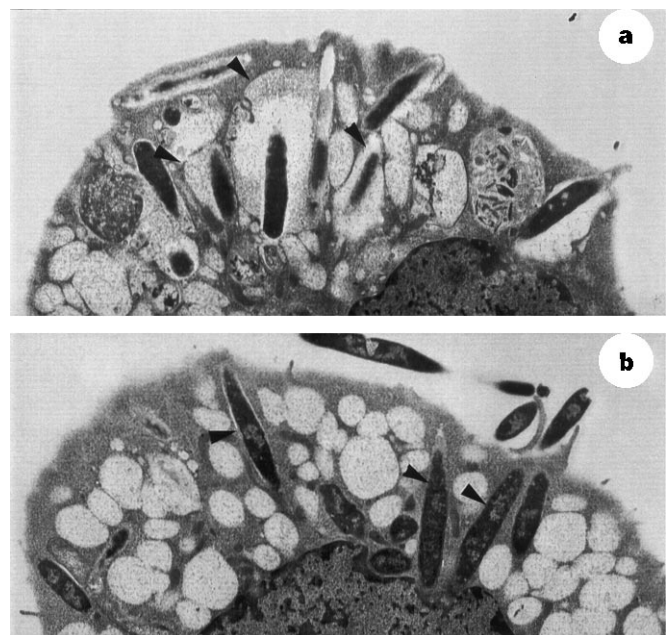


Figure 3 FimH-internalized *E. coli* are harboured in morphologically distinct compartments from opsonized *E. coli*. By 1 hour after internalization, *E. coli* phagocytosed via FimH are predominantly seen in tight-fitting vacuoles (arrowheads in b), whereas *E. coli* phagocytosed via opsonizing antibody are predominantly within membrane-bound vacuoles containing diffuse matrix material surrounding the bacterium (arrowheads in a).

Although the binding of FimH to CD48 is mannose-sensitive, this molecule is distinct from CD11/CD18 and CD66, which are surface molecules implicated in FimH recognition by granulocytes^{17,18}. CD48 is abundant on macrophages¹⁹ and is believed to contribute to antigen recognition by immune cells²⁰. In some cases, CD48 was transiently colocalized with non-opsonized FimH-expressing bacteria during phagocytosis (Fig. 4a, b), and antibody directed at CD48 in a dose-dependent way blocked the ingestion of non-opsonized FimH-expressing bacteria (Fig. 4g), implicating the GPI-linked CD48 as the putative FimH receptor in macrophages as well as mast cells. As many GPI-linked moieties are translocated intracellularly through a pathway distinct from the classical endosome-lysosome pathway involving the formation of glycolipid-enriched microdomains (a specialized subset of which comprise caveolae)²¹, we considered that non-opsonized *E. coli* might avoid the usual macrophage arsenal by using such a mechanism. Nystatin, filipin²² and methyl β -cyclodextrin²³, which disrupt lipid microdomains by cholesterol chelation, all specifically blocked

internalization of FimH-internalized bacteria but not opsonin-internalized bacteria (Fig. 4h, i). Caveolin, a frequent component of lipid-rich microdomains, co-localized with FimH-expressing bacteria during internalization (Fig. 4c–f). Comparison of the internalization of GPI-anchored proteins and of transmembrane proteins has indicated that the uptake of GPI-anchored molecules is different and is associated with glycolipid-enriched domains²⁴, but our results now provide evidence for a role in bacterial internalization.

We have attempted to resolve the prevailing paradox that *E. coli* and other enterobacteria express on their surface organelles of adhesion that promote bacterial colonization and infection but also facilitate uptake of the pathogen by phagocytes. We have shown that FimH-mediated binding to macrophages is actually advantageous to the bacterial population. We propose that by binding to the FimH-receptor CD48, non-opsonized *E. coli* gain access to a lipid-processing pathway that bypasses the normal phagocytic killing mechanisms. Only coming into play in cases of antibody deficiency,

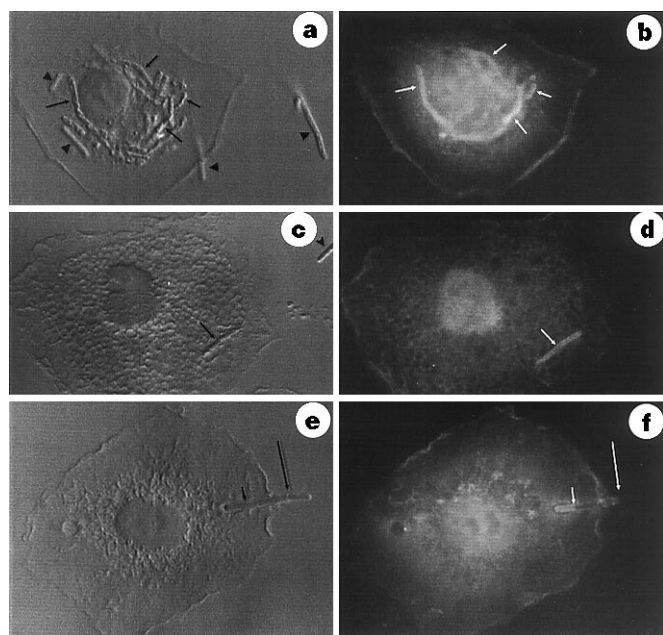
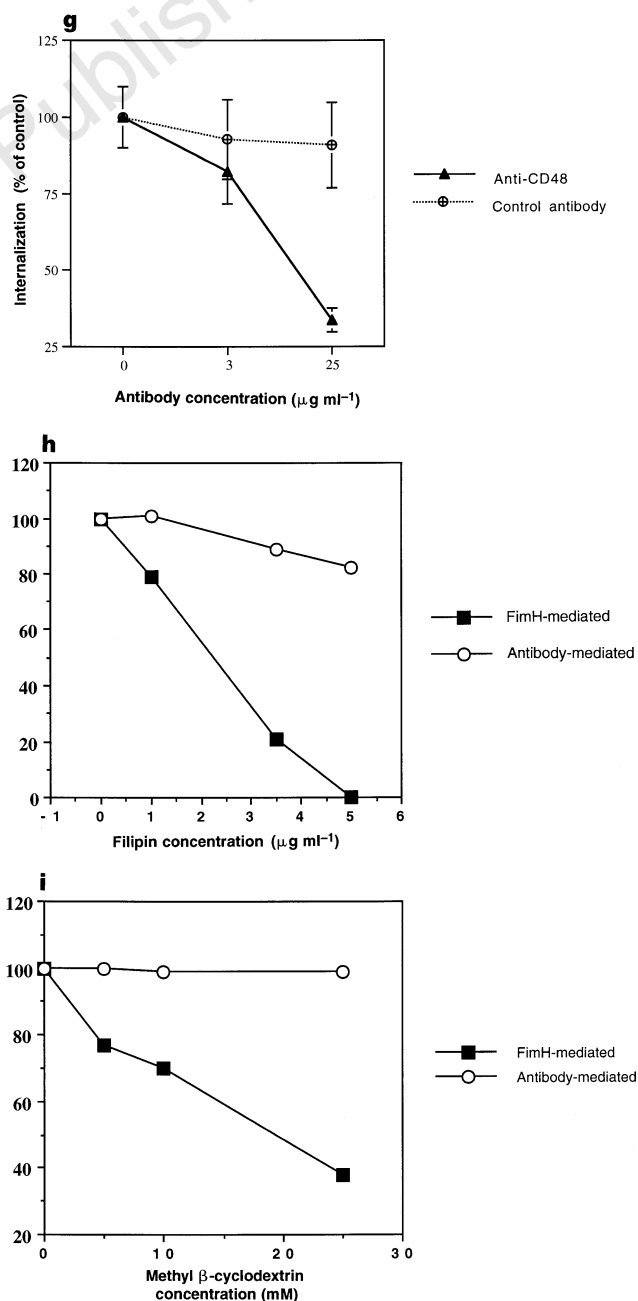


Figure 4 Use of CD48 and lipid trafficking pathways by FimH-internalized bacteria. CD48 immunofluorescence colocalizes with ingested bacteria (arrows in **b**), but not noningested bacteria (arrowheads in the corresponding differential interference contrast, DIC image, **a**). Caveolin immunofluorescence colocalizes with bacteria during internalization (arrow in **d**), but not unbound bacteria (arrowhead in DIC image, **c**). Caveolin immunofluorescence advancing along bacterial shaft (arrows in **f** and corresponding DIC image, **e**). **g**, Antibodies to CD48 inhibit FimH-mediated internalization, while control antibodies to the Fc γ RIII receptor have no effect. Filipin (**h**) and methyl β -cyclodextrin (**i**) specifically inhibit FimH-mediated, but not antibody-mediated, internalization of bacteria. Data on nystatin not shown.



the capacity of FimH-expressing enterobacteria to seek intracellular refuge may explain opportunistic infections by these organisms in opsonin-deficiency states and at body sites with poor opsonizing capacity (such as the bladder mucosa)¹⁻³. □

Methods

Macrophage culture. Preparation was from bone marrow of BALB/c mice as described²⁵, and verified as containing >99% macrophages by immunostaining with the monoclonal antibody F4/80. All experiments were done in HEPES α -medium in serum-free conditions.

Model system. The model for non-opsonic binding to macrophages was *E. coli* strain ORN103 expressing the plasmid pSH2, encoding the entire type-1 fimbrial gene cluster (including FimH)²⁶. The isogenic FimH⁻ mutant, *E. coli* ORN103(pUT2002)²⁷, does not bind; purely opsonic binding was induced by coating the FimH⁻ strain with specific antibody raised against *E. coli* ORN103 (pUT2002) in mice (1:600–1:200 dilution of heat-inactivated antibody). Preimmune sera did not induce binding, suggesting that binding of FimH⁻, opsonized bacterium to macrophages was mediated by antibody. The opsonization step had no effect on bacterial viability. The bacteria used for Fig. 1b were *E. coli* J96, a uropathogenic clinical isolate that naturally expresses type-1 fimbriae, including the FimH adhesin and Pile, its FimH⁻ isogenic derivative²⁸.

Intracellular viability assays. These were done by modification of methods previously used for other bacteria²⁹. Macrophage monolayers on coverslips were infected with bacteria at a multiplicity of infection (MOI) of 1 or less. After internalization, coverslips were treated with 100 $\mu\text{g ml}^{-1}$ gentamicin for 6 min–1 h (6 min kills >99% and 1 h kills all detectable extracellular bacteria for the 24-h time point). Intracellular *E. coli* were counted by plating onto MacConkey agar plates after solubilization of the infected macrophages in 0.1% Triton X100 in PBS. The viability stains (Fig. 1c, d) (from Molecular Probes) were used according to recommended protocols. For these experiments, the MOI was 100, infected macrophages were fixed with 4% paraformaldehyde and permeabilized for 30 s with 0.1% Triton X100.

Oxidative burst and acidification. The oxidative burst was detected by chemiluminescence⁵. For detection of intracellularly released oxygen metabolites, macrophages were loaded with 2 $\mu\text{g ml}^{-1}$ dihydrorhodamine(Sigma)³⁰ in Hanks balanced salt solution (HBSS) for 30 minutes before bacterial addition, and mean fluorescence per macrophage was followed for 10–22 individual macrophages per experiment using an excitation wavelength of 485nm. Images were captured under identical conditions and mean fluorescence for individual cells measured over time using IPLab Spectrum image processing software (Signal Analytics). The time course represents mean data from a representative experiment in which 16 individual macrophages were followed. The pH change of the bacterial compartment was detected by infecting macrophages with FITC-labelled *E. coli*, under either opsonic or non-opsonic conditions, and following the ratio of emission at excitation wavelengths of 490 nm/450 nm with time¹⁶.

Electron microscopy. Transmission electron microscopy was done on macrophages grown in 24-well tissue-culture plates. Following a 20-min incubation with bacteria, cultures were rinsed to remove non-adherent bacteria and then incubated for 1 hour before fixation with 2% glutaraldehyde. Specimens for electron microscopy were processed as described⁵.

Antibodies and immunofluorescence staining. Immunofluorescence was done by standard procedures. Briefly, FITC-labelled bacteria were added to macrophage cultures at an MOI of 50, and allowed to bind and interact with macrophages for 30 min. After paraformaldehyde fixation and brief solubilization in Triton, cells were stained with primary and secondary antibodies for 30 min each and mounted on slides with 50% glycerol in PBS. Rat anti-mouse CD48 antibodies were obtained from Serotec (Oxford), and used for staining at a dilution of 1:4. Secondary goat anti-rat IgG antibody labelled with rhodamine was from Zymed; rabbit anti-mouse caveolin antibodies were from Transduction Labs; secondary goat anti-rabbit IgG antibody labelled with rhodamine was from Boehringer Mannheim.

Internalization blocking experiments. Following a 60-min incubation with anti-CD48 antibody, bacteria were added for 20 min to allow phagocytosis to occur, gentamicin treatment was used to eliminate extracellular bacteria, and viable intracellular bacteria were determined. Nystatin (Sigma) was used at 0–

25 $\mu\text{g ml}^{-1}$, filipin (Sigma) was used at 0–5 $\mu\text{g ml}^{-1}$, and methyl β -cyclodextrin (Sigma) was used at 0–25 mM. Macrophage cultures were first pretreated for 30 min with the drug in serum-free medium, FITC-labelled bacteria were added in the presence of drug, allowed to bind for 15 min, rinsed, and bound bacteria were then allowed to internalize for 30 min, and fixed with 4% paraformaldehyde. The proportion of bound bacteria internalized in each condition was determined by labelling extracellular bacteria with rhodamine-conjugated antibody against *E. coli* raised in rabbit. This antibody does not penetrate the paraformaldehyde-fixed macrophage membrane and labels only extracellular bacteria (data not shown). As all bacteria are prelabelled with FITC, the proportion internalized under each condition is readily calculated.

Received 12 June; accepted 24 July 1997.

- Brauner, A., Kaijser, B. & Kuhn, I. Recurrent *Escherichia coli* bacteraemia—clinical characteristics and bacterial properties. *J. Infect.* **28**, 49–57 (1994).
- Seneca, H. Urinary-tract infections: etiology, microbiology, pathophysiology, diagnosis and management. *J. Am. Geriatr. Soc.* **29**, 359–369 (1981).
- Kreger, B. E., Craven, D. E., Carling, P. C. & McCabe, W. R. Gram-negative bacteremia. III. Reassessment of etiology, epidemiology and ecology in 612 patients. *Am. J. Med.* **68**, 332–343 (1980).
- Malaviya, R., Ikeda, T., Ross, E. & Abraham, S. N. Mast-cell modulation of neutrophil influx and bacterial clearance at sites of infection through TNF- α . *Nature* **381**, 77–80 (1996).
- Tewari, R. et al. Neutrophil activation by nascent FimH subunits of type 1 fimbriae purified from the periplasm of *Escherichia coli*. *J. Biol. Chem.* **268**, 3009–3015 (1993).
- Goetz, M. B. & Silverblatt, F. J. Phagolysosome formation by polymorphonuclear leukocytes after ingestion of *Escherichia coli* that express type 1 pili. *J. Infect. Dis.* **156**, 229–233 (1987).
- Keith, B. R., Harris, S. L., Russell, P. W. & Orndorff, P. E. Effect of type 1 piliation on *in vitro* killing of *Escherichia coli* by mouse peritoneal macrophages. *Infect. Immun.* **58**, 3448–3454 (1990).
- Gbarah, A. et al. *Shigella flexneri* transformants expressing type 1 (mannose-specific) fimbriae bind to, activate, and are killed by phagocytic cells. *Infect. Immun.* **61**, 1687–1693 (1993).
- Lock, R. et al. Neutrophil killing of two type 1 fimbriae-bearing *Escherichia coli* strains: dependence on respiratory burst activation. *Infect. Immun.* **58**, 37–42 (1990).
- Ofek, I., Goldhar, J., Keisari, Y. & Sharon, N. Nonopsonic phagocytosis of microorganisms. *Annu. Rev. Microbiol.* **49**, 239–276 (1995).
- May, A. K., Bloch, C. A., Sawyer, R. G., Spengler, M. D. & Pruett, T. L. Enhanced virulence of *Escherichia coli* bearing a site-targeted mutation in the major structural subunit of type 1 fimbriae. *Infect. Immun.* **61**, 1667–1673 (1993).
- Kisilevsky, P. V., Schwan, W. R., Amundsen, S. K., Duncan, J. L. & Schaeffer, A. J. *In vivo* expression and variation of *Escherichia coli* type 1 and P pili in the urine of adults with acute urinary tract infections. *Infect. Immun.* **57**, 1656–1662 (1989).
- Hagber, L. et al. Adhesion, hemagglutination, and virulence of *Escherichia coli* causing urinary tract infections. *Infect. Immun.* **31**, 564–570 (1981).
- Saukkonen, K., Cabellos, C., Burroughs, M., Prasad, S. & Tuomanen, E. Integrin-mediated localization of *Bordetella pertussis* within macrophages: role in pulmonary colonization. *J. Exp. Med.* **173**, 1143–1149 (1991).
- Sturgill-Koszycki, S. et al. Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase. *Science* **263**, 637–639 (1994).
- Eisenberg, L. G., Goldman, W. E. & Schlesinger, P. H. *Histoplasma capsulatum* modulates the acidification of phagolysosomes. *J. Exp. Med.* **177**, 1605–1611 (1993).
- Gbarah, A., Gahmberg, C. G., Ofek, I., Jacobi, U. & Sharon, N. Identification of the leukocyte adhesion molecules CD11 and CD18 as receptors for type 1-fimbriated (mannose-specific) *Escherichia coli*. *Infect. Immun.* **59**, 4524–4530 (1991).
- Sauter, S. L., Rutherford, S. M., Wagener, C., Shively, J. E. & Hefta, S. A. Identification of the specific oligosaccharide sites recognized by type 1 fimbriae from *Escherichia coli* on nonspecific cross-reacting antigen, a D66 cluster granulocyte glycoprotein. *J. Biol. Chem.* **268**, 15510–15516 (1993).
- Wong, Y. W., Williams, A. F., Kingsmore, S. F. & Seldin, M. F. Structure, expression, and genetic linkage of the mouse BCM1 (OX45 or Blast-1) antigen. Evidence for genetic duplication giving rise to the BCM1 region on mouse chromosome 1 and the CD2/LFA3 region on mouse chromosome 3. *J. Exp. Med.* **171**, 2115–2130 (1990).
- Davis, S. J. & van der Merwe, P. A. Structure and ligand interactions of CD2: implications for T-cell function. *Immunol. Today* **17**, 177–187 (1996).
- Parton, R. G. Caveolae and caveolins. *Curr. Opin. Cell Biol.* **8**, 542–548 (1996).
- Schnitzer, J. E., Oh, P., Pinney, E. & Allard, J. Filipin-sensitive caveolae-mediated transport in endothelium: reduced transcytosis, scavenger endocytosis, and capillary permeability of select macromolecules. *J. Cell Biol.* **127**, 1217–1232 (1994).
- Yancey, P. G. et al. Cellular cholesterol efflux mediated by cyclodextrins. Demonstration of kinetic pools and mechanism of efflux. *J. Biol. Chem.* **271**, 16026–16034 (1996).
- Keller, G. A., Siegel, M. W. & Caras, J. W. Endocytosis of glycopospholipid-anchored and transmembrane forms of D4 by different endocytic pathways. *EMBO J.* **11**, 863–874 (1992).
- Clohisy, D. R., Bar-Shavit, Z., Chappel, J. & Teitelbaum, S. L. 1,25-Dihydroxyvitamin D3 modulates bone marrow macrophage precursor proliferation and differentiation. *J. Biol. Chem.* **262**, 15922–15929 (1987).
- Orndorff, P. E. & Falkow, S. Organization and expression of genes responsible for type 1 piliation in *Escherichia coli*. *J. Bact.* **159**, 736–744 (1984).
- Abraham, S. N., Sun, D., Dale, J. B. & Beachey, E. H. Conservation of the D-mannose-adhesion protein among type 1 fimbriated members of the family *Enterobacteriaceae*. *Nature* **336**, 682–684 (1988).
- Keith, B. R., Maurer, L., Spears, P. A. & Orndorff, P. E. Receptor-binding function of type 1 pili effects bladder colonization by a clinical isolate of *Escherichia coli*. *Infect. Immun.* **53**, 693–696 (1986).
- Berger, K. H. & Isberg, R. R. Intracellular survival by *Legionella*. *Meth. Cell Biol.* **45**, 247–259 (1994).
- Blue, J. L., Gallois, A., Schneider, J. C., Parini, J. P. & Tschirhart, E. A double-labeling fluorescent assay for concomitant measurements of $[\text{Ca}^{2+}]$ and O_2 production in human macrophages. *Biochim. Biophys. Acta* **1244**, 79–84 (1995).

Acknowledgements. We thank D. Haslam and I. Ofek for discussion, R. Henry for electron microscopy, and P. Orndorff for the gift of *E. coli* J96 and its FimH derivative. This work was supported in part by research grants from the NIH and from an award from the Jewish Hospital Foundation.

Correspondence and requests for materials should be addressed to S.N.A. (e-mail: abraham@pathbox.wustl.edu).

Mitotic chromatin regulates phosphorylation of Stathmin/Op18

Søren S. L. Andersen*, Anthony J. Ashford, Régis Tournebise, Olivier Gavet†, André Sobel†, Anthony A. Hyman‡ & Eric Karsenti‡

EMBL, Cell Biology Programme, Meyerhofstr. 1, D-69117 Heidelberg, Germany

†INSERM U440, 17, rue du Fer à Moulin, F-75005 Paris, France

Meiotic and mitotic spindles are required for the even segregation of duplicated chromosomes to the two daughter cells. The mechanism of spindle assembly is not fully understood, but two have been proposed that are not mutually exclusive^{1–3}. The 'search and capture' model suggests that dynamic microtubules become progressively captured and stabilized by the kinetochores on chromosomes, leading to spindle assembly^{3,4}. The 'local stabilization' model proposes that chromosomes change the state of the cytoplasm around them, making it more favourable to microtubule polymerization^{2,5–9}. It has been shown^{10,11} that Stathmin/Op18 inhibits microtubule polymerization *in vitro* by interaction with tubulin¹², and that overexpression in tissue culture cells of non-phosphorylatable mutants of Stathmin/Op18 prevents the assembly of mitotic spindles¹³. We have used *Xenopus* egg extracts and magnetic chromatin beads¹⁴ to show that mitotic chromatin induces phosphorylation of Stathmin/Op18. We have also shown that Stathmin/Op18 is one of the factors regulated by mitotic chromatin that governs preferential microtubule growth around chromosomes during spindle assembly.

It has recently been shown that Stathmin/Op18 is involved in the regulation of microtubule dynamics and spindle assembly^{12,13}. We therefore decided to examine the regulation of its phosphorylation during the transition from interphase to metaphase in *Xenopus* egg extracts. The analysis of immunoprecipitates of Stathmin/Op18 from interphase and mitotic egg extracts by SDS–PAGE (polyacrylamide gel electrophoresis) revealed a triplet (Fig. 1a). Two-dimensional gel electrophoresis of the same immunoprecipitates revealed four spots (Figs 1b, 2b). We could not detect any difference between interphase and mitotic immunoprecipitates using either of these techniques. Treatment of the immunoprecipitates with alkaline phosphatase reduced the triplet to one band using SDS–PAGE (data not shown), and the four spots to two using two-dimensional gel electrophoresis, indicating that both interphase and mitotic extracts contain two partly phosphorylated Stathmin/Op18 isoforms, referred to as n1, n2, p1, p2 (refs 10, 15) (Fig. 1b). However, there was no significant incorporation of phosphate into interphasic or mitotic Stathmin/Op18, indicating a low phosphate turnover rate (Fig. 1a). Taken together, these results indicate that Stathmin/Op18 is not differentially phosphorylated between interphase and mitosis in *Xenopus* egg extracts.

This was surprising given a previous report showing that Stathmin/Op18 is hyperphosphorylated in mitotic cells¹³. We wondered whether the lack of difference in Stathmin/Op18 phosphorylation between interphase and mitotic extracts was due to the absence of nuclei in these extracts. To test this idea we added increasing amounts of chromatin beads from interphase or mitosis to interphase and mitotic extracts. The addition of interphase bead nuclei to an interphase or to a mitotic extract did not result in an increase in Stathmin/Op18 phosphorylation (Fig. 1c and data not shown). However, two additional bands that incorporated ³²P appeared on

SDS–PAGE of immunoprecipitates following the addition of mitotic chromatin beads to a mitotic extract (Fig. 1c). Hyperphosphorylation of Stathmin/Op18 was dose dependent and reached a maximum at 50,000–100,000 mitotic chromatin beads per 1-μl extract, which corresponds to 5,000–10,000 nuclei per 1-μl extract (Fig. 1d, e).

We did not detect any significant incorporation of phosphate into Stathmin/Op18 in mitotic extracts lacking chromosomes (n1, n2, p1, p2; Fig. 1), so we suspected that the chromatin beads induced the hyperphosphorylation of Stathmin/Op18 by changing the balance between the activity of phosphatases and kinases in the extract. To investigate the role of phosphatases in determining the phosphorylation state of Stathmin/Op18, we added phosphatase

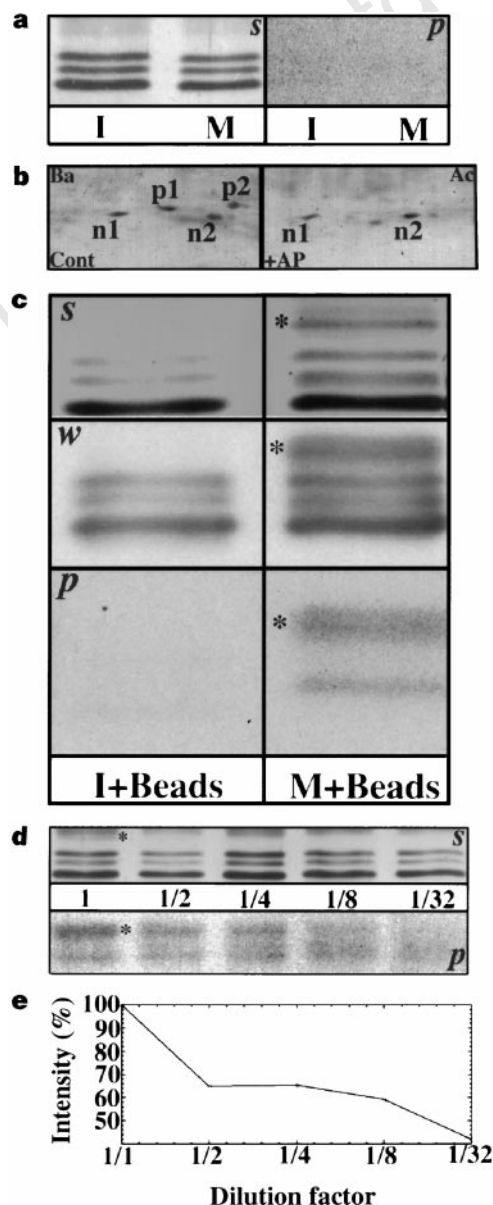


Figure 1 Hyperphosphorylation of Stathmin/Op18 during mitosis requires mitotic chromatin. **a**, [γ -³²P]ATP-labelled Stathmin/Op18 from interphase (I) and mitotic (M) extracts; s, silver staining; p, autoradiography, 15% SDS–PAGE. **b**, Two-dimensional gels of purified mitotic Stathmin/Op18 treated without (Cont) or with alkaline phosphatase (+AP) reveals n1, n2, p1, p2; n, non-phosphorylated; p, phosphorylated; Ba, basic; Ac, acidic. **c**, Stathmin/Op18 hyperphosphorylation and two high-molecular-weight forms (asterisks) during mitosis (M + Beads), but not interphase (I + Beads); w, western blot. **d**, Mitotic chromatin beads resuspended at 50,000–100,000 beads per 1 μl extract (1), diluted (dilution factor) with extract, and then [γ -³²P]ATP-labelling for 15 min. **e**, Quantification of d.

* Present address: Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544-1014, USA.

‡ These authors contributed equally to this work.

inhibitors to interphase and mitotic extracts. Inhibition of type 2A phosphatases (PP2A)¹⁶ with 0.5 μ M okadaic acid in interphase extracts had no effect on the phosphorylation of Stathmin/Op18 (data not shown). However, in mitotic extracts, this resulted in an SDS-PAGE phosphorylation pattern identical to that produced by adding mitotic chromatin (Fig. 2a). Total inhibition of phosphatase 1 with 3 μ M inhibitor 2 had no significant effect on Stathmin/Op18 phosphorylation¹⁷. By two-dimensional gel electrophoresis, Stathmin/Op18 resolved as six spots after treatment with okadaic acid or mitotic chromatin (Fig. 2b). The clearest difference between control extracts and extracts treated with mitotic chromatin or okadaic acid was the appearance of two extra phosphorylated Stathmin/Op18 spots, termed p3 and p4, which probably arise from the phosphorylation of p1 and p2 (Fig. 2b). This result could mean that Stathmin/Op18 is actively dephosphorylated by PP2A, or that Stathmin/Op18 becomes phosphorylated because inhibition of PP2A activates a kinase that phosphorylates Stathmin/Op18, or the mechanism may be more complex. To test these possibilities, we added *in vivo* phosphorylated Stathmin/Op18 to mitotic extracts, and observed that 0.5 μ M okadaic acid blocked dephosphorylation (Fig. 2c). When this experiment was repeated with chromatin beads instead of okadaic acid, we did not detect a significant block of Stathmin/Op18 dephosphorylation (data not shown). However, we

measured¹⁸ an approximately 10% reduction in total PP2A activity in the presence of high concentrations of mitotic chromatin beads, but did not detect any change in the Cdc2 kinase activity (data not shown). Thus the mechanism by which chromatin induces Stathmin/Op18 phosphorylation remains unclear.

In order to investigate which of the three *Xenopus* Stathmin/Op18 (ref. 15) phosphorylation sites were actively phosphorylated in the extract, we mutated all three serines to alanines in all possible combinations, and added endogenous amounts of the recombinant proteins to interphase and mitotic extracts in the presence of [γ -³²P]ATP (Fig. 3). No incorporation of phosphate was observed when all three serines were mutated (Δ all), showing that Ser 16 (PKA site), Ser 25 (MAPK site) and Ser 39 (Cdc2 site)¹⁵ can account for the phosphorylation patterns observed in both interphase and mitotic extracts. Addition of mutants with only one functional site to interphase and mitotic extracts revealed that only Ser 39 incorporates phosphate (Fig. 3). However, in interphase, both a functional Ser 16 and a functional Ser 39 are required for a wild-type phosphorylation level of Stathmin/Op18, whereas Ser 25 appears dispensable (Fig. 3). In mitosis, either Ser 16 or Ser 25, together with Ser 39, is required to obtain wild-type phosphorylation levels (Fig. 3, quantification not shown). These results indicate that the Cdc2 site (Ser 39) is the principal site phosphorylated in mitosis, and that there is some degree of cooperativity between the phosphorylation of Ser 39 and Ser 16/Ser 25, as previously proposed¹⁹.

The previous results indicated that the activity of Stathmin/Op18 could be regulated by phosphorylation. To address this question we tested the effect on spindle stability of mutant Stathmin/Op18 containing all phosphorylation sites changed to alanines. The addition of this mutant to preassembled spindles resulted in a reduction of microtubule mass, decreasing spindle size significantly (by 20%) compared with those after the addition of an equivalent amount of wild-type Stathmin/Op18 (Fig. 4A, B; graphs are based on measurement of 2,320 spindles, and the s.d. of each point was 5–7 μ m). At mutant protein concentrations above 200 μ g ml⁻¹, spindle stability was strongly affected, causing a collapse of most of the spindles (the endogenous Stathmin/Op18 concentration is 5–6 μ M, or 100 μ g ml⁻¹; data not shown and ref. 12). We determined that the observed spindle shortening occurs within 3 min of Stathmin/Op18 addition (Fig. 4B), which is in the range of microtubule half-life in spindles assembled *in vitro*²⁰. Addition of Stathmin/Op18 during spindle assembly had the same effect (data not shown). These data indicate that phosphorylation of Stathmin/Op18 is required to turn off its inhibitory effect on microtubule polymerization¹³. Because adding Stathmin/Op18 results in spindle shortening we thought that removing Stathmin/Op18 might result

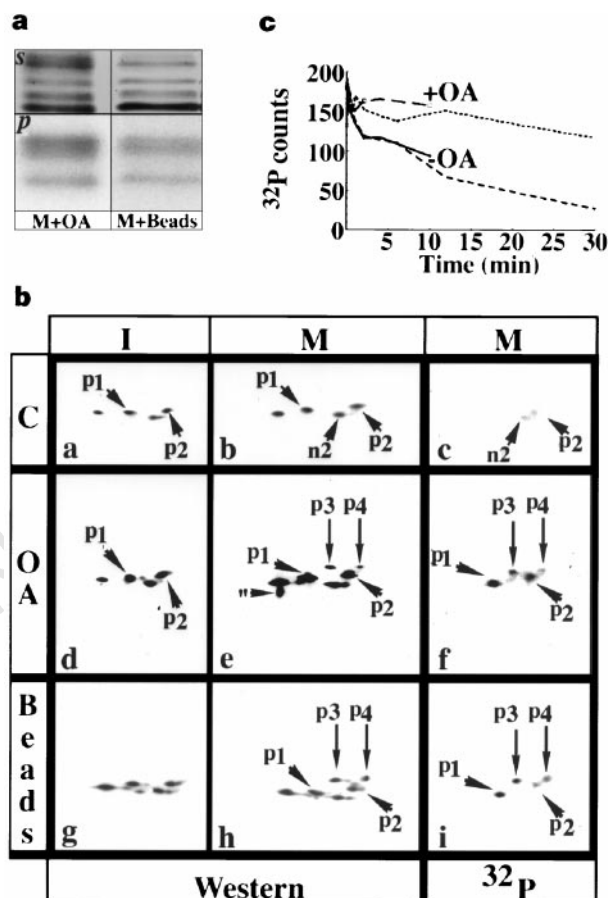


Figure 2 a type 2A phosphatase regulates Stathmin/Op18 phosphorylation. **a**, Stathmin/Op18 from mitotic extracts treated with chromatin (M+Beads) or 0.5 μ M okadaic acid (M+OA); s, silver staining; p, autoradiography, 15% SDS-PAGE. **b**, Two-dimensional (right, acidic end) phosphorylation pattern of Stathmin/Op18; control (c), okadaic acid (OA) and chromatin bead (Beads) treatment in interphase (I) or mitotic (M) extracts; p3 and p4 only appear after treatment with OA or mitotic chromatin beads, also required for ³²P-labelling of p1 and p2. Note the more complex pattern with OA than mitotic chromatin, notably a non-reproducible spot (II); Western, western blotting; ³²P, autoradiography. **c**, Dephosphorylation of Stathmin/Op18 is blocked by 0.5 μ M okadaic acid in mitotic extracts.

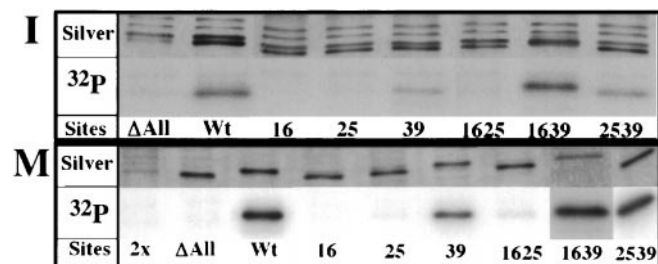


Figure 3 Stathmin/Op18 is phosphorylated mainly on the Cdc2 site in mitosis. Constructs of Stathmin/Op18 were added to an interphase (I) and a Stathmin/Op18-depleted mitotic extract (M) for 15 min together with 1 μ Ci [γ -³²P]ATP per 1- μ l extract, and then immunoprecipitated (15% SDS-PAGE; Silver, silver staining; ³²P, autoradiography). See text for sites. 2x indicates double depletion without addition of recombinant protein. Addition of proteins to a non-depleted mitotic extract gave similar results. In interphase, phosphorylation on the Cdc2 site could be due to a kinase other than Cdc2. In the presence of 0.5 μ M okadaic acid, no specific site showed a significant increase in ³²P incorporation, but p3 and p4 forms were observed.

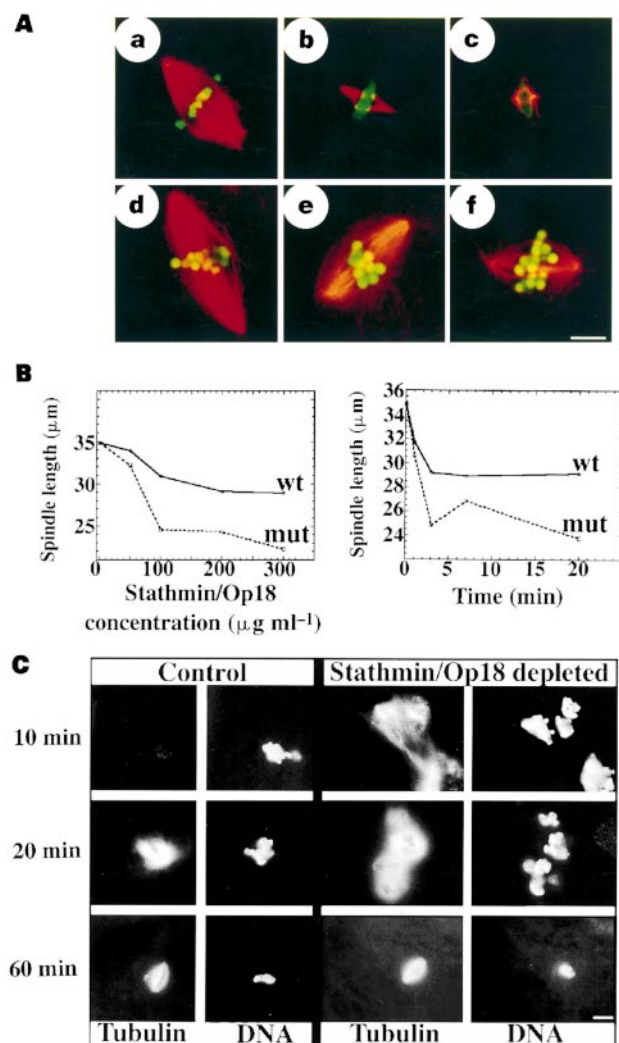


Figure 4 Mutant Stathmin/Op18 causes rapid shortening of the mitotic spindle. **A, B**, Pre-formed mitotic sperm and bead spindles were incubated with 100 (**a, d**), 200 (**b, e**) and 300 (**c, f**) $\mu\text{g ml}^{-1}$ mutant Stathmin/Op18, fixed after 20 min, and the mean spindle length was measured (DNA, green; microtubules, red); scale bar, 5 μm . **B**, Quantification of the sperm spindles shown in **A**; using wild-type (wt) and mutant (mut) Stathmin/Op18. Left, length; right, time, for which 100 $\mu\text{g ml}^{-1}$ Stathmin/Op18 was used. **C**, Immunoprecipitation of Stathmin/Op18 results in accelerated microtubule nucleation around mitotic chromatin. Mitotic chromatin was added to control or Stathmin/Op18-depleted extracts and time points taken. Addition of recombinant wild-type *Xenopus* Stathmin/Op18 to *in vivo* concentrations rescued the effect of immunodepletion. Scale bar, 25 μm .

in promotion of microtubule assembly and nucleation. Immunodepletion¹² of 90–95% of Stathmin/Op18 from the extract had a significant effect during the early stages of spindle assembly, when large sperm microtubule asters were observed (Fig. 4C), as previously described¹². Moreover, we observed an acceleration in microtubule formation around mitotic chromatin beads (approximately 10–20 min; Fig. 4C). However, at later stages of bead and sperm spindle assembly there were no differences between the Stathmin/Op18 depleted and non-depleted extracts (Fig. 4C), and direct measurements of microtubule dynamics have shown that Stathmin/Op18 is not the principal regulator of microtubule catastrophes in extracts¹⁷ or *in vitro*³⁰. These data suggest that, although lowering the concentration of Stathmin/Op18 affects microtubule dynamics in the extract, as well as the kinetics of the early stages of spindle assembly⁸, microtubule dynamics are regulated by other factors, still allowing spindle assembly. One candidate molecule is XKCM1, a *Xenopus* kinesin-like protein that increases the frequency of catastrophes and is required for spindle assembly²¹. XKCM1 and other factors may make microtubules sufficiently dynamic to allow assembly of mitotic spindles in the absence of Stathmin/Op18.

The most striking observation is that mitotic chromatin causes a dose-dependent increase in the phosphorylation level of Stathmin/Op18. Although we do not understand exactly how this is achieved, our data strongly suggest that a type 2A phosphatase (PP2A)¹⁶ and

Cdc2 kinase are involved. Indeed, when PP2A is inhibited in mitotic extracts, phosphorylation of Stathmin/Op18 probably occurs through Cdc2 kinase, as the Cdc2 site is the main site phosphorylated, and inhibition of PP2A in interphase extracts does not result in hyperphosphorylation (data not shown). We can think of two possible mechanisms to explain how mitotic chromatin could induce the phosphorylation of Stathmin/Op18 in mitotic extracts. The simplest mechanism would involve an activity present on chromosomes that blocks PP2A, which dephosphorylates Stathmin/Op18. But this seems to be in contradiction with the lack of inhibition of Stathmin/Op18 dephosphorylation by chromatin beads. The other possibility is that a kinase²² present on chromosomes (perhaps activated by partial PP2A inhibition) is responsible for the phosphorylation of Stathmin/Op18. This mechanism could explain why mitotic chromatin does not significantly block Stathmin/Op18 dephosphorylation in the extract. Further work will be needed to determine the details of this pathway. The progressive increase in Stathmin/Op18 phosphorylation, which parallels the number of beads added to the extracts, suggests that there is a gradient of Stathmin/Op18 phosphorylation/activity around chromosomes. This may be part of a general mechanism involved in regulating preferential microtubule growth around chromosomes during spindle assembly. Chromatin may also regulate the phosphorylation of Stathmin/Op18 during mitosis in other systems¹³. Moreover, in the absence of chromosomes, endogenous Stathmin/

Op18 is phosphorylated to the same extent in both interphase and mitosis (Fig. 1a), suggesting that its main function is not to regulate the global change in microtubule dynamics that occurs during the transition from interphase to mitosis^{11–13}. Rather, we think this change is caused by global phosphorylation of microtubule-associated proteins, which reduces their affinity for microtubules^{23–25}, allowing destabilizing factors to make microtubules more dynamic.

We predict that other molecules involved in the regulation of microtubule assembly are regulated in a similar manner to Stathmin/Op18. Good candidate molecules include spindle-associated microtubule-associated proteins^{23–25}, as well as the spindle-associated motor proteins XKCM1, XKLP1 and CENP-E^{21,26,27}. □

Methods

Extracts and spindles. *Xenopus* egg extracts (10,000g) arrested in the second meiotic metaphase (CSF extracts; H1 kinase activity, 15–30 pmol $\mu\text{L}^{-1}\text{min}^{-1}$; ref. 28) for sperm, magnetic bead spindle and magnetic interphase bead nuclei (50% competent for nuclear transport) assembly were prepared fresh and cycled as described^{14,29}. For interphase extracts (H1 kinase activity, 2–4 pmol $\mu\text{L}^{-1}\text{min}^{-1}$), eggs were activated for 15 min in MMR/4, 0.2 $\mu\text{g mL}^{-1}$ A23187 (Sigma), 200 $\mu\text{g mL}^{-1}$ cycloheximide. When spindles had assembled (Fig. 4A), 1/10 volume control buffer, wild-type or mutant Stathmin/Op18 diluted in CSF-XB was added. Purified wild-type and mutant (serines changed to alanines at positions 16, 25, 38 and 63) human Stathmin/Op18 (from A. Sobel) were in PBS at 10 mg mL^{-1} .

Immunoprecipitations. Extracts 5–40 μL were labelled with 1 μCi [γ -³²P]ATP per μL extract for 15 min at 20 °C. For the immunoprecipitation, 2–10 volumes cold SBMI (100 mM NaF, 80 mM β -glycerophosphate, 20 mM $\text{Na}_2\text{P}_2\text{O}_7$, 20 mM EDTA, pH 7.2, with 1 μM microcystin-LR (Gibco BRL), 1 mM PMSF and 10 $\mu\text{g mL}^{-1}$ each of leupeptin, pepstatin and aprotinin) containing Affi-prep protein A beads (Biorad) with rabbit anti-Stathmin/Op18 antibodies precoupled (partial gift from T. J. Mitchison¹²) was added for 30 min at 4 °C, followed by two washes with 200 μL SBMI, four washes with 200 μL PBS plus 100 mM NaCl, 0.1% Tx-100, protease inhibitor mix. When present, magnetic beads were removed after SBMI addition but before the immunoprecipitation. In the presence of okadaic acid, $\Delta 90\text{CyclinB}$ was added to the CSF extracts to prevent inactivation of Cdc2 kinase. Interphase bead nuclei added for <5 min, fresh DNA beads, or polyglutamic acid control beads added to a mitotic or interphase extract had no effect on phosphorylation (Fig. 1, data not shown).

Statistics. Comparison of spindle length in the presence of wild-type and mutant Stathmin/Op18; A χ^2 -test verified that spindle lengths follow a normal distribution, and Bartlett's test showed that the standard deviations are equal. The statistical test subsequently consisted of a *t*-test, $t = (\mu_1 - \mu_2) / [s(1/n_1 + 1/n_2)^{1/2}]$. $P > 95\%$ ($P < 0.05$) is considered statistically significant.

***Xenopus* Stathmin/Op18 cloning and expressions.** The 5' and 3' oligonucleotides used for PCR cloning of *Xenopus* Stathmin/Op18 (XLXO35)¹⁵ contained a *Nde*I and a *Hind*III site, respectively: GGGAATTCC ATATGTGTGACTCTGATATTAAGTAAACAGC and CCAAGCTTCTATC ACTTCTCAGAAGTTCTTTCATTC. A *Xenopus* cDNA library was prepared using purified polyadenylated RNA (RNeasy, Qiagen) from XL177 cells and a cDNA synthesis kit (Boehringer-Mannheim). *Xenopus* Stathmin/Op18 was subcloned into pMW172. Mutagenesis of *Xenopus* Stathmin/Op18 was performed according to the Quickchange site-direct mutagenesis kit protocol (Stratagene). All constructs were sequenced in pMW172 from the T7 promoter (EMBL sequencing service). *Xenopus* Stathmin/Op18 was expressed in *Escherichia coli* (BL21DE3), induced at 0.3–0.5 absorbance at 600 nm (0.1 mg mL^{-1} IPTG) for 3 h at 37 °C.

Purifications. Pellets from 1–2 l of cells were sonicated on ice for 3 $\times$ 20 s in 50–100 ml 20 mM Tris-HCl pH 8.3, 1 mM EGTA, 1 mM EDTA, 0.1% β -mercaptoethanol, and protease inhibitor mix. This was centrifuged at 30,000g for 10 min at 4 °C, and batch-absorbed to 10–20 ml DEAE (Pharmacia) in 20 mM Tris, pH 7, 200 mM NaCl for 20 min at 4 °C. DEAE was removed (4,000g, 10 min, 4 °C), the supernatant filtered through Kleenex paper, and ammonium sulphate powder was added to 60% for 30 min at 4 °C. The $(\text{NH}_4)_2\text{SO}_4$ pellet was recovered by centrifugation at 15,000g, 10 min, 4 °C, then resuspended in 5–10 ml 30 mM K-MES, pH 5.7, 0.1% β -mercaptoethanol and protease

inhibitor mix, and dialysed overnight at 4 °C against 400 volumes of the same buffer. The dialysate was centrifuged at 14,000g, 10 min, 4 °C, loaded onto a Mono-S PC 1.6/5 SMART column (Pharmacia) equilibrated with buffer A (20 mM K-MES, pH 5.7). Stathmin/Op18 was eluted with a step-gradient in 100 μL at 25% buffer B (buffer A with 500 mM NaCl). The yield never exceeded 200 μg per litre of cells (BCA; Pierce).

Dephosphorylation. Stathmin/Op18 was 100-fold purified from 2 $\times$ 600 μL okadaic acid-treated [γ -³²P]ATP-labelled mitotic extract; 5 volumes SBMI was added after 15 min labelling, and the mix was heated at 95 °C for 5 min, centrifuged at 100,000g, 10 min, 2 °C, followed by 5-fold concentration in Centricon-30, desalting into CSF-XB on a 5-ml Kwiksep column (Pierce), and 2-fold concentration in Centricon-30. This concentrate was mixed at a maximum dilution ratio of 1:1 with extracts and samples taken into sample buffer (Fig. 2c).

Miscellaneous. Pure Stathmin/Op18 (Fig. 1b) was obtained by immunoprecipitation, elution for 1 h at 4 °C with 200 $\mu\text{g mL}^{-1}$ anti-Stathmin/Op18 peptide¹² in XB, dilution with 10 volumes XB, concentration on Centricon-10 and incubation with 2 units alkaline phosphatase μL^{-1} for 1 h at 37 °C.

Received 7 May; accepted 7 July 1997.

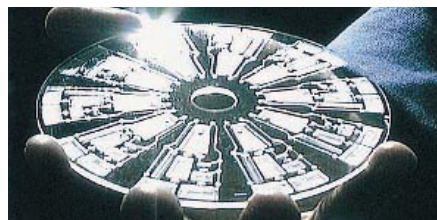
- Waters, J. C. & Salmon, E. D. Pathways of spindle assembly. *Curr. Opin. Cell Biol.* **9**, 37–43 (1997).
- Hyman, A. A. & Karsenti, E. Morphogenetic properties of microtubules and mitotic spindle assembly. *Cell* **84**, 401–411 (1996).
- Nicklas, R. B. How cells get the right chromosomes. *Science* **275**, 632–637 (1997).
- Kirschner, M. W. & Mitchison, T. J. Beyond self assembly: from microtubules to morphogenesis. *Cell* **45**, 329–342 (1986).
- Marek, L. E. Control of spindle form and function in Grasshopper spermatocytes. *Chromosoma* **68**, 367–398 (1978).
- Karsenti, E., Newport, J. & Kirschner, M. Respective roles of centrosomes and chromatin in the conversion of microtubule arrays from interphase to metaphase. *J. Cell Biol.* **99**, 475–575 (1984).
- Nicklas, R. B. & Gordon, G. W. The total length of spindle microtubules depends on the number of chromosomes present. *J. Cell Biol.* **100**, 1–7 (1985).
- Zhang, D. & Nicklas, R. B. The impact of chromosomes and centrosomes on spindle assembly as observed in living cells. *J. Cell Biol.* **129**, 1287–1300 (1995).
- Zhang, D. & Nicklas, R. B. Chromosomes initiate spindle assembly upon experimental dissolution of the nuclear envelope in grasshopper spermatocytes. *J. Cell Biol.* **131**, 1125–1131 (1995).
- Sobel, A. Stathmin: a relay phosphoprotein for multiple signal transduction? *Trends Biochem. Sci.* **16**, 301–305 (1991).
- Belmont, L., Mitchison, T. & Deacon, H. W. Catastrophic revelations about Op18/stathmin. *Trends Biochem. Sci.* **21**, 197–198 (1996).
- Belmont, L. D. & Mitchison, T. J. Identification of a protein that interacts with tubulin dimers and increases the catastrophe rate of microtubules. *Cell* **84**, 623–631 (1996).
- Marklund, U., Larsson, N., Gradin, H. M., Brattsand, G. & Gullberg, M. Oncoprotein 18 is a phosphorylation-responsive regulator of microtubule dynamics. *EMBO J.* **15**, 5290–5298 (1996).
- Heald, R. et al. Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* **382**, 420–425 (1996).
- Maucuer, A., Moreau, J., Mechali, M. & Sobel, A. Stathmin gene family: phylogenetic conservation and developmental regulation in *Xenopus*. *J. Biol. Chem.* **268**, 16420–16429 (1993).
- Mayer-Jackel, R. E. & Hemmings, B. A. Protein phosphatase 2A — a 'ménage à trois'. *Trends Cell Biol.* **4**, 287–291 (1994).
- Tournebise, R. et al. Distinct roles of PP1 and PP2A-like phosphatases in control of microtubule dynamics during mitosis. *EMBO J.* (in the press).
- Felix, M. A., Cohen, P. & Karsenti, E. Cdc2 H1 kinase is negatively regulated by a type 2A phosphatase in the *Xenopus* early embryonic cell cycle: evidence from the effects of okadaic acid. *EMBO J.* **9**, 675–683 (1990).
- Larsson, N., Melander, H., Marklund, U., Osterman, O. & Gullberg, M. G2/M transition requires multisite phosphorylation of oncoprotein 18 by two distinct protein kinase systems. *J. Biol. Chem.* **270**, 14175–14183 (1995).
- Sawin, K. E. & Mitchison, T. J. Poleward microtubule flux in mitotic spindles assembled *in vitro*. *J. Cell Biol.* **112**, 941–954 (1991).
- Walczak, C. E., Mitchison, T. J. & Desai, A. XKCM1: a *Xenopus* kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. *Cell* **84**, 37–47 (1996).
- Samanta, A. & Greene, M. I. A kinase associated with chromatin that can be activated by ligand-p185^{Neu} or epidermal growth factor-receptor interactions. *Proc. Natl Acad. Sci. USA* **92**, 6582–6586 (1995).
- Ookata, K. et al. Cyclin B interaction with microtubule-associated protein 4 (MAP4) targets p34Cdc2 kinase to microtubules and is a potential regulator of M-phase microtubule dynamics. *J. Cell Biol.* **128**, 849–862 (1995).
- Masson, D. & Kreis, T. E. Binding of E-MAP-115 to microtubules is regulated by cell cycle-dependent phosphorylation. *J. Cell Biol.* **131**, 1015–1024 (1995).
- Andersen, S. S. L., Buendia, B., Domínguez, J. E., Sawyer, A. & Karsenti, E. Effect on microtubule dynamics of XMAP230, a microtubule-associated protein present in *Xenopus laevis* eggs and dividing cells. *J. Cell Biol.* **127**, 1289–1299 (1994).
- Vernos, I. et al. Xklp1, a chromosomal *Xenopus* kinesin-like protein essential for spindle organization and chromosome positioning. *Cell* **81**, 117–127 (1995).
- Lombillo, V. A., Nislow, C., Yen, T. J., Gelfand, V. I. & McIntosh, J. R. Antibodies to the kinesin motor domain and CENP-E inhibit microtubule depolymerization-dependent motion of chromosomes *in vitro*. *J. Cell Biol.* **128**, 107–115 (1995).
- Felix, M. A., Clarke, P., Coleman, J., Verde, F. & Karsenti, E. In *The Cell Cycle: A Practical Approach* (eds Fantes, P. & Brooks, R.) 253–283 (IRL, New York, 1994).
- Shamu, C. E. & Murray, A. W. Sister chromatid separation in frog egg extracts requires DNA topoisomerase II activity during anaphase. *J. Cell Biol.* **117**, 921–934 (1992).
- Curmi, P. A. et al. The Stathmin tubulin interaction *in vitro*. *J. Biol. Chem.* (in the press).

Acknowledgements. We thank D. Chrétien, K. Deigaard, J. Domínguez, R. Heald, J. Howard, T. J. Mitchison, I. T. Möst, A. Nebreda, M. Way, T. Wittmann and M. Zerial for discussions and reagents.

Correspondence and requests for materials should be addressed to S.S.L.A. (e-mail: andersen@embl-heidelberg.de).

Tools for biomedical research labs

Offering up a variety of devices and reagents for biomedical sciences, this edition includes a single-cell superfusion system, a bioluminescence assay for monitoring cell-cell interactions, and a cell sizer and counter.



Gamera Bioscience's LabCD technology.

LabCD

From Gamera Bioscience

While this product still has to undergo beta testing during the next year, the concept is interesting enough for a preview of what should be available in 1999

Based on disposable compact discs and CD-player technology, this system is designed for diagnostic testing and other key laboratory procedures, such as DNA amplification and detection, cell counting and identification, and immunoassays. The manufacturer hopes that this technology will be used not only in the laboratory, but also by untrained users in clinical, non-clinical and home-testing environments. The use of CD technology was developed as an answer to the problem of pumping and valving very small volumes of fluid. Gamera intends this device to provide portable, automated cell counting and identification and be the first system to perform DNA amplification and detection in one step. The user simply places a drop of blood, saliva or urine in an opening in the centre of the disc, which has been altered to contain sample input ports, internal capillaries and reagent reservoirs. When the disc is spun, samples are moved via centrifugal force, the 'player' analyses the processed fluids and provides a quantitative result.

Reader Enquiry No. 100

Transgenic animals

From Taconic

A new 32-page book entitled, Access to Transgenic Technologies — An Introduction to the Taconic Transgenic Exchange

This booklet is now available to assist researchers creating and using transgenic animals. It provides a straightforward, comprehensive outline of transgenic rodent technology. Beginning with a brief history of transgenics, the book includes a thorough review of the considerations for maintaining, sharing and distributing transgenic research animals. The book provides both scientists and Offices of

Technology Transfer with current information about transgenic models: maintenance concerns, relevant intellectual property and licensing. Taconic is an international supplier of murine pathogen-free (MPF) laboratory animals, including the *mdr1a* knockout mouse and over a dozen others.

Reader Enquiry No. 101

DAD-12 superfusion system

From ALA Scientific Instruments

A turnkey instrument for focal fluid application

This superfusion system is designed to deliver a known concentration of a drug or substance to a single cell or tissue under study (exchange time: 10–15 ms). PC-based software controls the system through a parallel-port interface, allowing the user to programme time-based sequences that turn on/off valves and apply pressure to drive up to 12 solutions out of the micromanifold, which uses a 100- μ m i.d. tip. Cell(s) under study are completely bathed in the applied solution ('sewer pipe' effect). This system includes a valve/reservoir manifold, micro-bore Teflon tubing, a quartz micromanifold, an interface, software and an optional laptop controller. The system can be customized, as well.

Reader Enquiry No. 102

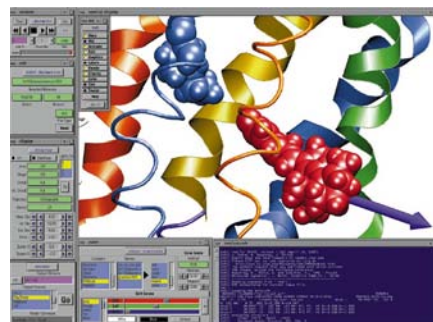
Freeware

VMD molecular visualization

From the Theoretical Biophysics Group, Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign

A free molecular visualization program designed for interactive image display

The software is capable of animation, manipulation and analysis of proteins, nucleic acids, lipid membranes and other biological molecules. It supports OpenGL and SGI GL with stereo displays.



VMD, a free molecular visualization program.

Moreover, the CAVE and input formats for many popular renderers are included (VRML, POV-Ray and Raster3D). VMD can directly read Protein Data Bank (PDB) files, X-PLOR PSF/DCD files and Amber CRD files, and it is able to use Babel to convert from other file formats. The wide variety of methods for rendering and colouring molecules offered by VMD includes VdW, licorice, cartoon and surface models. Animation information can come from a trajectory file or by connecting to a running simulation program. The VMD scripting language is based on Tcl/Tk and may be used to query molecular information, perform analysis and to display the results interactively, yielding a powerful base for implementing new methods in structure visualization and analysis. To further method development, the complete C++ source code is freely available, along with documentation for using and modifying the program. Version 1.2 of VMD runs under most versions of Unix, including IRIX, Linux and Solaris, and will soon be ported to Windows NT. See the VMD home page for more information: <www.ks.uiuc.edu/Research/vmd/>

Reader Enquiry No. 103

Reagents and equipment

Enzyme substrate catalogue

From Glycosynth

This catalogue offers a series of substrates for the clinical biochemist, microbiologist, genetic and plant researcher, and diagnostic testing laboratory

The company will fill custom orders of enzyme substrates, such as Rose-gal (6-chloro-3-indolyl β -D-galactopyraside). These materials are used not only for routine assays and diagnostic tests but also in trying to identify inherited diseases with a metabolic base and in the development of specific diagnostic tests. In addition, the manufacturer offers AMC compounds of high purity, such as AMC-L-alanine.

Reader Enquiry No. 104

OptiFreeze

From ICN

Dimethyl sulphoxide (DMSO) cryopreservation medium for optimal recovery of preserved cells

This medium has been formulated for improved freezing results through the use of an efficient cryopreservant. It is said to minimize the destruction and damage

normally associated with freezing and storing cells, and provides improved recovery rates of preserved cells. OptiFreeze contains 10 per cent dimethyl sulphoxide, 10 per cent heat-inactivated FBS (from USDA-certified abattoirs in the US) and is balanced with Iscove's modified Dulbecco's medium with HEPES. It comes ready for use with no dilution necessary and each lot of medium is sterile filtered through a capsule filter of 0.2 μm . Endotoxin is $<1 \text{ ng ml}^{-1}$, as determined by *Limulus* amoebocyte lysate chromogenic assay and is negative for mycoplasma infection. The manufacturer states that this medium is stable for at least five freeze-thaw cycles and for three years at -20°C . It is suitable for use in serum-dependent cell cultures. ICN also offers Cellvation cryopreservation medium for use in serum-free cell cultures.

Reader Enquiry No. 105

Human lung (mast cell) tryptase

From Cortex Biochem

Intended for in vitro and manufacturing use only, the enzyme is a tetramer with four active enzymatic sites

Tryptase is stored in granules inside mast cells and is released by exocytosis when the mast cell is activated. Tryptase has been implicated in the pathophysiology of allergic asthma. It is expected that the availability of this product will be an important tool for researchers who are developing new treatments for a wide variety of inflammatory allergic responses.

Reader Enquiry No. 106

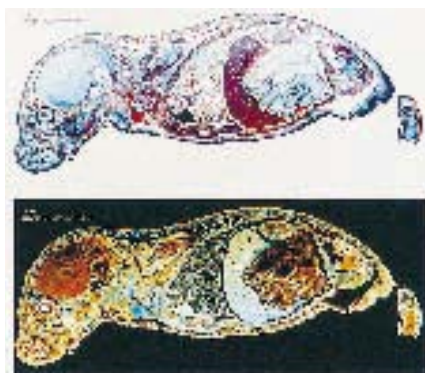
BioLite

From Packard Instrument

A new luminescence assay for monitoring live cell interactions and virus infectivity

This assay uses a non-toxic luminescence reagent for labelling cell membranes or bacterial cell walls, as well as for use in the study of cell-cell interactions and for measuring mammalian and bacterial cells. Most recently, the assay has been used to measure the release of viruses from infected, labelled cells. As the virus buds from the cell, it carries with it a coating of the cell membrane containing the BioLite label. Not only can the amount of virus released be measured, but the subsequent uptake of the virus by uninfected, unlabelled cells can also be measured. BioLite is the newest member in a family of ConstantQuanta reagents that have long-lived luminescent signals: LucLite, for measuring the luciferase reporter gene for transcriptional control assays, and CytoLite, for measuring cell viability (proliferation and toxicity).

Reader Enquiry No. 107



PathScan Enabler for imaging whole slides.

PathScan Enabler

From Bilaney Consultants

Digital imaging for whole slides

This system has been designed to produce high-resolution images of whole-slide cross sections, in colour and with high clarity. Where the pathology of the entire sample is important, a conventional microscope suffers from a restricted field of view, making it difficult to understand the macro characteristics of an image. PathScan Enabler overcomes this problem by scanning in the microscope slide at a resolution of 2,700 d.p.i., producing an image that can be viewed, measured, manipulated and analysed on a PC screen. For comparison, a conventional CCD camera resolves at 300 d.p.i. Areas of interest can then be studied, further magnified (to about $\times 30$), dropped into documents or presentations, or shared with colleagues over the Internet.

Reader Enquiry No. 108

Z2 cell counter and sizer

From Coulter

A compact, mercury-free cell count and size analyser

The Z2 reports a cell size distribution in mean cell size and mean cell volume (MCV). Statistics are reported either over the entire

size distribution or over a portion of the distribution as defined by moveable cursors. Typical applications include cell counts for animal blood analysis, tissue-culture studies, oncology research, algae studies, and research with spermatozoa and yeast.

Reader Enquiry No. 109

MACS-VA500

From Don Whitley Scientific

A variable atmosphere workstation that has a controlled environment to allow the study and isolation of microaerophilic organisms such as Campylobacter and Helicobacter spp

The technical advancements of this workstation supplies up to four gases and, for maximum flexibility, the operator can pre-select varying ratios of hydrogen (0–9 per cent), carbon dioxide (0–20 per cent), and oxygen (0–20 per cent), with the balance of the atmosphere being nitrogen. The atmosphere within the chamber is microprocessor-controlled for maximum accuracy — atmospheric requirements are entered through a keypad/display unit, which also generates system information and error messages. The MACS-VA500 offers a simple and cost-effective way for laboratories to culture atmospherically demanding organisms. Like the MG500 workstation, this unit has a plate capacity of 540 plates. The operator can select the intensity of high-level illumination required, and an automated control system maintains the relative humidity of the environment while also removing excess condensation from the chamber before being evaporated. The VA500 also offers the convenience of the porthole transfer system, which allows sample transfer and serves as armholes for the operator (the oval shape and bare hands facility are said to give improved user comfort and flexibility). Each porthole can be used to transfer ten 90-mm plates at the same time as an operator's arms are inserted or withdrawn from the cabinet.

Reader Enquiry No. 110

ADVERTISEMENTS

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT

PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray
Phone: +44 191 222 8612
Fax: +44 191 222 8129
Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 114

Gramsch Laboratories

New: Endomorphin -1 & -2
Custom Peptide Synthesis:
Shipped within one week:
10 AA DM 400.00; 15 AA DM 520.00;
20 AA DM 660.00; 10-30 mg unpurified;
buy only when satisfied with the HPLC profile

Antiserum Production Service:
you prepare the recombinant protein (in paa-gel or soluble), we produce the antiserum

Distributor for Japan: FUNAKOSHI CO, Tokio 113
Gramsch Laboratories
D-85247 SCHWABHAUSEN, Germany
Phone +49-8138-1769 or +49-8131-5679-0
Fax +49-8138-1260 or +49-8131-5679-24

READER ENQUIRY NO. 95

Grosslab Junior

From Shandon Lipshaw

A bench-top workstation for gross examination and sectioning of pathology specimens

Constructed from high-quality stainless steel, the workstation contains a large sink (46 cm × 15 cm × 10 cm) with rapid positive drainage. It is supplied with a vacuum breaker-protected water tap, which prevents pollutants from entering the potable water supply. There is abundant fluorescent light directed to the work surface and, to minimize contamination, the Grosslab features a hands-free control that activates water with an infrared signal (this activation signal can be manually overridden, if necessary). Each unit comes complete with a magnetic instrument holder, integral centimetre rulers, a towel rack and liquid dispenser, as well as a strainer for small specimens and curettings. For convenient access, appliances can be plugged into the Grosslab Junior directly. According to the manufacturer, hazardous fumes are removed quietly and efficiently, either through a duct adapter connected to the facility exhaust system or by passage through special formaldehyde vapour filters. The filter drawers are located in the front of the unit for easy access.

Reader Enquiry No. 111

Immunology-based systems

CamVir-1MAB

Stains From BioGenex

A monoclonal antibody for staining paraffin-embedded tissues infected with papillomavirus

This antibody stains certain types of human papillomavirus (HPV), including types 16, 18, 31, 33 and 35, which have been associated as a major risk factor in the subsequent development of cervical cancer. The high-risk HPV 16 group is almost equally associated with cervical intraepithelial neoplasia (CIN) and cervical cancer: monoclonal antibody CamVir-1 can be used to stain formalin-fixed, paraffin-embedded tissue sections for the presence of papillomavirus type 16. This antibody is available in concentrated form or as a ready-to-use reagent optimized for use with the 'super sensitive' biotin-streptavidin detection systems, and is currently available for research use only.

Reader Enquiry No. 112

ProCount

From Becton Dickinson

Immunocytometry Systems

A progenitor cell enumeration system

Designed to be a fully integrated flow cytometry system for CD34+ cell enumeration in studies of stem cell transplantation, this system combines new software with



BDIS's ProCount flow cytometry system.

reagents that are formulated to reduce variability in CD34 cell counting. An accurate measure of the CD34 cell count is necessary for dose requirement protocols in studies of stem cell transplantation. ProCount reagents are under US regulatory review for use with peripheral blood, mobilized peripheral blood and leukopheresis products. The system software has been optimized for use on FACSCalibur, FACS Sort and FACScan flow cytometers from Becton Dickinson. However, samples prepared with ProCount reagents may be acquired and analysed manually on other brands of flow cytometers. ProCount software automates instrument set up and control, sample acquisition, sample analysis, and data management. In addition, the system provides clear reports containing valuable, easy-to-read information for patient management. All results, including CD34+ cells in the harvest, may be exported to a tab-delimited file.

Reader Enquiry No. 113

Phosphoprotein antibody sampler pack

From Zymed Laboratories

Highly specific antibodies for phosphoserine, phosphothreonine and phosphotyrosine

Reversible protein phosphorylation plays an important role in numerous biochemical pathways and functions to alter one or more properties of a protein, including activity, association (complex formation), conformation, localization and/or stability. In eukaryotic cells, serine, threonine and tyrosine represent the major substrates for phosphorylation by protein kinases. Traditional methods for detecting protein phosphorylation relied upon the use of radioactive ^{32}P , which is both hazardous and laborious. Zymed now offers this sampler pack, which contains highly specific polyclonal antibodies to phosphoserine, phosphothreonine and phosphotyrosine. These antibodies provide a safe and simple way to monitor the changes in the phosphorylation status of a given protein. Each antibody has been put through a cascade of antigen affinity purifications to provide improved specificity. The sample pack contains 20- μg aliquots of each antibody, which are useful for western blotting, immunoprecipitation, ELISA and tissue staining.

Reader Enquiry No. 114

Human IL-5 ELISA kit

From Wako

An interleukin-5 ELISA kit that offers high specificity and $\pm 1 \text{ pg ml}^{-1}$ detection limits

The sensitivity of this sandwich-based assay is stated to be greater than the several hundred nanogram per millilitre detection limits that are characteristic of classical bioassays. It uses a monoclonal capture antibody and polyclonal secondary antibody, which does not demonstrate crossreactivity with murine IL-5 or other human cytokines, says Wako. The 96-assay kit is useful for research that is concerned with eosinophil and B-cell differentiation, and subsequent antibody release mechanisms, including various autoimmune implications.

Reader Enquiry No. 115

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

LIPOSOMES

NEW

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

P.O. Box 650126, D-81215 Munich
For USA call 1-800 Liposome

READER ENQUIRY NO. 62

2-Hybrid System Media

Over 400 recipes
and formulae in stock

CONVENIENCE

CONSISTENCY

ECONOMY

RELIABILITY

Contact BIO 101 or your local distributor for ordering & information
(800) 424-6101 ♦ fax: (760) 598-0116
web: www.bio101.com ♦ e-mail: info@bio101.com

READER ENQUIRY NO. 12